

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS SOCIAIS E HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO**

**PREVISÃO DO MERCADO ACIONÁRIO POR MEIO
DE REDES NEURAIS MLP E REDES NEURAIS
KOHONEN EM PERÍODO DE CRISE ECONÔMICA**

DISSERTAÇÃO DE MESTRADO

André Pacheco Miranda

**Santa Maria, RS, Brasil
2013**

**PREVISÃO DO MERCADO ACIONÁRIO POR MEIO DE
REDES NEURAIS MLP E REDES NEURAIS KOHONEN EM
PERÍODO DE CRISE ECONÔMICA**

André Pacheco Miranda

Dissertação de mestrado apresentado ao Curso de Mestrado do Programa de Pós-Graduação em Administração, Área de Sistemas e Finanças, da Universidade Federal de Santa Maria (UFSM, RS), como requisito final para obtenção do grau de **Mestre em Administração**.

Orientador: Prof. Dr. Luis Felipe Dias Lopes

**Santa Maria, RS, Brasil
2013**

Ficha catalográfica elaborada através do Programa de Geração Automática da
Biblioteca Central da UFSM, com os dados fornecidos pelo(a) autor(a):

Miranda, André Pacheco
PREVISÃO DO MERCADO ACIONÁRIO POR MEIO DE REDES NEURAIIS
MLP KOHONEN EM PERÍODOS DE CRISE ECONÔMICA / André Pacheco
Miranda.-
2013.
106 f.; 30cm

Orientador: Luis Felipe Dias Lopes
Dissertação (mestrado) - Universidade Federal de Santa
Maria, Centro de Ciências Sociais e Humanas, Programa de
Pós-graduação em Administração, RS, 2013.

1. Redes Neurais 2. Crises econômicas 3. Análise
técnica I. Dias Lopes, Luis Felipe II. Título.

© 2013

Todos os direitos autorais reservados a André Pacheco Miranda. A reprodução de partes ou do
todo deste trabalho só poderá ser com autorização por escrito do autor.

Endereço: Rua Venâncio Aires, n. 576, Bairro Centro, Santa Maria, RS, 97010-000

Fone (0xx)55 99699393; End. Eletr: andmirandapippi@gmail.com

**Universidade Federal de Santa Maria
Centro de Ciências Sociais e Humanas
Programa de Pós-Graduação em Administração**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**PREVISÃO DO MERCADO ACIONÁRIO POR MEIO DE REDES
NEURAS MLP E REDES NEURAS KOHONEN EM PERÍODO DE
CRISE ECONÔMICA**

elaborado por
André Pacheco Miranda

Como requisito final para obtenção do grau de
Mestre em Administração

COMISSÃO EXAMINADORA:

Luis Felipe Dias Lopes, Dr. (UFSM)
(Presidente/Orientador)

Viviane Cátia Köhler, Dra. (UFSM)
(Primeiro examinador)

Paulo Sérgio Ceretta, Dr. (UFSM)
(Segundo examinador)

Santa Maria, 28 de agosto de 2013.

“...Saudações aos loucos. Aos desajustados. Aos rebeldes. Aos criadores de caso. Às peças redondas nos buracos quadrados. Aos que veem as coisas de forma diferente. Eles não gostam de regras. E eles não têm nenhum respeito pelo status quo. Você pode citá-los, discordar deles, glorificá-los ou difamá-los. Mas a única coisa que você não pode fazer é ignorá-los. Porque eles mudam as coisas. Eles empurram a raça humana para frente. Enquanto alguns os veem como loucos, nós os vemos como gênios. Porque as pessoas que são loucas o suficiente para achar que podem mudar o mundo, são as que de fato, mudam...”

(Steve Jobs)

“...Note aquelas coisas que não podem ser vistas, tomando consciência daquilo que não é óbvio...”

(Miyamoto Musashi)

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Administração
Universidade Federal de Santa Maria

PREVISÃO DO MERCADO ACIONÁRIO POR MEIO DE REDES NEURAI MLP E REDES NEURAI KOHONEN EM PERÍODO DE CRISE ECONÔMICA

AUTOR: ANDRÉ PACHECO MIRANDA

ORIENTADOR: LUIS FELIPE DIAS LOPES

Data e Local da Defesa: Santa Maria, 28 de agosto de 2013.

As oscilações no mercado acionário por meio de crises econômicas, riscos de deflação e armadilhas de liquidez são pontos críticos na análise de risco, que ocasionam discrepâncias na execução de um determinado escopo no mercado de renda variável. A crise da inadimplência do subprime em 2007/2008 que obteve uma das maiores repercussões nos mercados financeiros fundou novas discussões em relação ao controle de risco na tomada de decisão do investidor. No mercado acionário, a análise de risco busca auxiliar o investidor na tomada de decisões, para isso utiliza-se de ferramentas e métodos estatísticos para tentar prever os movimentos do mercado. Com base nestas afirmações anteriores e com o intuito de auxiliar o investidor na tomada de decisão mediante a uma crise econômica, este trabalho, do tipo exploratório, objetivou-se desenvolver e treinar duas redes neurais com aprendizados diferenciados sem o problema da “caixa preta” dos métodos a fim de comparar quais das duas tem melhor previsão em períodos de crises econômicas. Como variáveis de entrada para as redes neurais utilizou-se o retorno do volume semanais do Ibovespa no período de 12/08/2002 até 30/05/2011 e um setup desenvolvido a partir da Teoria das Ondas de Elliott. Ou seja, estas duas redes neurais foram desenvolvidas, treinadas e validadas para antever os movimentos do mercado quando este apresentar oscilações provenientes de uma crise econômica. Como mencionado anteriormente para validar o estudo, foi comparado o poder de explicação dos dois métodos, antes de um ponto de uma provável crise. Conclui-se, portanto, que a analogia feita para a criação da teoria das ondas de Elliott entre a teoria do comportamento psicológico das massas e a seqüência de Fibonacci se mostrou incapaz de prever oscilações do mercado em uma série correspondente a uma crise econômica. Conclui-se, também, que redes neurais com aprendizados não supervisionados que utilizam variáveis temporais como variáveis de entrada tem a previsão superior no treinamento, mas inferiores na etapa mais determinante a validação das redes.

Palavras-chaves: Redes neurais. Crises econômicas. Análise técnica .

ABSTRACT

Dissertação de Mestrado
Programa de Pós-Graduação em Administração
Universidade Federal de Santa Maria

PREVISÃO DO MERCADO ACIONÁRIO POR MEIO DE REDES NEURAIAS MLP E REDES NEURAIAS KOHONEN EM PERÍODO DE CRISE ECONÔMICA

AUTHOR: ANDRÉ PACHECO MIRANDA

ADVISOR: LUIS FELIPE DIAS LOPES

Date and Place of Presentation: Santa Maria, august 28th, 2013.

Fluctuations in the stock market through economic crises, risks of deflation and liquidity traps are critical in the analysis of risk, which cause discrepancies in the execution of a particular scope in the equities market. The crisis in subprime insolvency in 2007/2008 which had a major impact on financial markets founded further discussions in relation to risk control in the decision making of investors. In the stock market risk analysis seeks to assist the investor in making decisions, for it makes use of statistical methods and tools to try to predict market movements. Based on these and previous statements in order to assist investors in making decisions through an economic crisis, this is an exploratory study aimed to develop and train two neural networks with differentiated learning without the problem of "black box" methods to compare which of the two has better forecast in periods of economic crisis. As input variables for the neural networks used the return of the volume of weekly Ibovespa in the period 12/08/2002 to 30/05/2011 and a setup developed from the Elliott Wave Theory. That is, these two neural networks were developed, trained and validated to predict market movements when it presents oscillations from an economic crisis. As mentioned earlier to validate the study compared the power of explanation of two methods before a point of probable attack. We conclude, therefore, that the analogy made for the creation of the theory of Elliott wave theory of psychological behavior of the masses and the Fibonacci sequence proved unable to provide for oscillations of the market in a series corresponding to an economic crisis. It was concluded, too, that neural networks with unsupervised learning using temporal variables as input variables has a higher prediction in training, but lower than most crucial step in the validation of systems.

Keywords: Neural networks. Economic crisis. Technical analysis.

LISTA DE TABELAS

Tabela 1- Função lógica “E”	29
Tabela 2- Função lógica “OU”	30
Tabela 3- Função lógica OU exclusivo “XOR”	31
Tabela 4 - Dados do treinamento da tabela de classificação de animais	44
Tabela 5 - Dados da tabela de classificação de animais	46
Tabela 6 - Entradas Saída Desejada	73
Tabela 7 - Entradas Decodificadas para as Redes Neurais	74
Tabela 8 - Taxa de Aprendizado	81
Tabela 9 - Tabela da Topologia da MLP	81
Tabela 10 - Erro Desejado	82
Tabela 11 - Validação Cruzada Rede Neural Kohonen	84
Tabela 12 - MLP VS Kohonen	87
Tabela 13 - Teste da “Caixa Preta”	89

LISTA DE FIGURAS

Figura 1 - Neurônio Matemático	25
Figura 2 - Gráfico da função lógica “E”	30
Figura 3 - Gráfico da função lógica “OU”	31
Figura 4 - Gráfico da função lógica OU exclusivo “XOR”	32
Figura 5 - Topologia de uma Rede Kohonen	41
Figura 6 - Fluxograma rede Kohonen	43
Figura 7 - Gráfico do Ponto de Parada Ideal	50
Figura 8 - Gráfico do Ponto de Afastamento	51
Figura 9 - Topologia Rede Neural Artificial MLP	53
Figura 10 - Ciclo do Mercado de Ações acompanhado das Ondas de Elliott	55
Figura 11 - Exemplos da terceira regra da Teoria das Ondas de Elliott	58
Figura 12 - Extensão das Ondas de Impulso	59
Figura 13 - Contagem da onda quatro, oitava regra da teoria das ondas	60
Figura 14 - Triângulo da nona regra da Teoria das Ondas de Elliott	61
Figura 15 - Etapas da Manipulação dos Dados	64
Figura 16 - Entradas da Rede Neural Perceptron de Múltiplas Camadas	66
Figura 17 - Entradas da Rede Neural Kohonen	67
Figura 18 - Aplicação da terceira regra de Elliott	69
Figura 19 - Aplicação da quinta regra de Elliott	70
Figura 20 - Aplicação da sexta regra de Elliott	70
Figura 21 - Contagem da onda quatro, oitava regra da teoria das ondas	71
Figura 22 - Triângulo da nona regra da Teoria das Ondas de Elliott	72
Figura 23 - Série retorno do volume semanal normalizada com raiz unitária	80
Figura 24 - Treinamento VS Validação Cruzada	83
Figura 25 - Série de Dados da Verificação	86

LISTA DE QUADROS

Quadro 1 - Funções de Transferência 26

LISTA DE ANEXOS

Anexo A – Classificação dos dados de entrada conforme os padrões encontrados na Teoria das Ondas de Elliott.....	103
Anexo B – Gráfico Resultado do Treinamento	104

SUMÁRIO

1 INTRODUÇÃO E MOTIVAÇÃO	13
1.1 Delimitando o problema de pesquisa.....	15
1.2 Objetivos.....	16
1.3 Justificativa	17
1.5 Etapas do Trabalho.....	18
2 FUNDAMENTAÇÃO TEÓRICA	19
2.1 Ibovespa.....	19
2.2 Crises no Mercado Acionário	19
2.2.1 Crise da Inadimplência do <i>Subprime</i> 2008.....	21
2.3 Breve Historia a Respeito das Redes Neurais Artificiais	22
2.4 Neurônio Biológico.....	23
2.5 Neurônio Matemático	24
2.5.1 Função de Ativação	25
2.5.2 Função de Transferência Regra de Propagação	26
2.5.3 Redes Neurais <i>Perceptron</i>	28
2.5.4 <i>Perceptron</i> Múltiplas Camadas (MLP)	32
2.5.4.1 Definição de uma MLP.....	33
2.5.4.2 Retropropagação do Erro com Aprendizagem Supervisionada	33
2.5.4.2.1 Propagação progressiva do sinal (Fase Forward).....	34
2.5.4.2.2 Reajuste dos Pesos (Fase Backward)	35
2.5.5 Redes Neurais Kohonen	37
2.5.5.1 Aprendizado Concorrente Processo Competitivo	38
2.5.5.2 Aprendizado Concorrente Processo Cooperativo	39
2.5.5.3 Aprendizado Concorrente Processo Adaptativo	40
2.5.5.4 Funcionamento Externo da Rede Neural Kohonen	41
2.5.5.5 Retropropagação do Erro com Aprendizagem Não Supervisionada.....	43
2.5.6 Treinamento / Aprendizagem, Validação Cruzada e Verificação das Redes Neurais	48
2.5.6.1 O Treinamento / Aprendizado.....	48
2.5.6.2 A Validação Cruzada.....	49
2.5.6.3 A Verificação	51
2.5.6.4 A Verificação Entre os Métodos	52
2.5.7 Topologias das Redes Neurais	52
2.5.8 Iluminando a “Caixa Preta”	53
2.6 A Teoria das ondas de Elliott.....	54
2.6.1 A Base Matemática da Teoria.....	57
3 MATERIAL E MÉTODOS	62
3.1 Manipulação dos Dados	62
3.2 Métodos de Pré-treinamento Rede Neural MLP	65
3.3 Métodos de Pré-treinamento Rede Neural Kohonen.....	66

3.4 Métodos de Treinamento	67
3.4.1 Entradas: Saída Desejada	68
3.4.1.1 Saída Desejada: 1ª Etapa	69
3.4.1.2 Saída Desejada: 2ª Etapa	72
3.4.1.3 Entrada das Redes Neurais: 3ª Etapa.....	74
3.4.2 Treinamento / Aprendizado Rede Neural <i>Perceptron</i>	74
3.4.3 Métodos da Validação Cruzada e Verificação.....	75
3.5 Testes utilizados	76
4 ANÁLISE E DISCUSSÕES DOS RESULTADOS	79
4.1 Rede neural MLP Treinamento e Validação Cruzada	79
4.2 Rede neural Kohonen Treinamento e Validação Cruzada	83
4.3 Verificação	85
4.4 Testes Realizados	88
4.5 Limitações	89
5 CONSIDERAÇÕES FINAIS	91
5.1 Sugestões para Trabalhos Futuros	93
REFERÊNCIAS	95

1 INTRODUÇÃO E MOTIVAÇÃO

Uma boa exploração do mercado acionário com o intuito de minimizar riscos sempre foi ambicionada por muitos investidores deste mercado a fim de buscarem vários métodos matemáticos e computacionais para controle de risco e entender a imprevisibilidade de títulos financeiros (HISSA, 2008). O gerenciamento de riscos em um projeto financeiro está intrinsecamente ligado aos conceitos de oportunidade e ameaças, que provem, principalmente, do acontecimento de fatores externos ao projeto. Entretanto, existe um fator externo denominado crise econômica que pode ser caracterizado como ainda mais imprevisível que o dia a dia do mercado financeiro.

O tema crises econômicas desperta um grande interesse por parte dos pesquisadores das áreas financeiras, pois estes fenômenos surgem no cenário econômico com certa frequência, atingindo economias, desestruturando países em um âmbito desigual, com espantosas reações emocionais (MELLO e SPOLADOR, 2007). As crises econômicas em um âmbito mundial tendem a transferir as incertezas aos seus respectivos mercados (HUYGHEBAERT e WANG, 2010). Além das turbulências econômicas, as crises trazem um colapso de insensatez aos preços das bolsas de valores, que, por sua vez, exercem uma pressão emocional especulativa negativa nos investidores.

Os surtos de euforia e pessimismo trazidos pelas pressões emocionais negativas em razão das crises econômicas levam a volatilidade das bolsas de valores ao redor do mundo a uma inconstância geral nos preços das ações. Essas volubilidades são responsáveis pelos grandes volumes nas posições vendidas dos ativos no mercado de renda variável (FARHI, 2006).

Para Hissa (2008), a análise de mercado não pode antecipar minuciosamente o futuro dos ativos ou de uma carteira de ações por depender da determinação de milhões de pessoas. Porém, as decisões tomadas com base nas pressões emocionais especulativas deixam o mercado um pouco menos eficiente, abrindo espaço para modelos de previsão. O uso de métodos e modelos matemáticos como forma de antever as possíveis movimentações do mercado financeiro não é uma prática nova para os pesquisadores das áreas de Administração, Economia e Inteligência Artificial como Kimoto (1990), Steiner (1995), Lawrence (1997), Sobreiro et al. (2009) e Oliveira et al. (2010). Esses sistemas para a previsão do mercado acionário vão desde as mais simples técnicas estatísticas, presentes em análises técnicas (ROBERT, 2007), lógica *fuzzy*, algoritmos genéticos (LAWRENCE, 1997; WUERGES e

BORBA, 2010), modelos de Markov (HASSAN, 2005), modelos econométricos (OLIVEIRA et al., 2010; COELHO et al., 2008), redes neurais probabilísticas com menor desempenho (KHASHEI et al., 2012), dentre outros (LAWRENCE, 1997; WUERGES e BORBA, 2010).

Entretanto, destas ferramentas citadas anteriormente, um dos modelos matemáticos menos utilizados para a previsão são as redes neurais, como pode ser visto em (LAWRENCE, 1997; KIM, 2004; HASSAN, 2005; PHUA, 2000; SMITH e GUPTA, 2000; WUERGES e BORBA, 2010, CUNHA et al., 2010). As principais dificuldades no uso das redes neurais como forma de antever as possíveis movimentações do mercado financeiro resultam da pouca familiaridade com elas entre os pesquisadores das áreas de ciências humanas. Isto ocorre por se tratar de métodos considerados mais modernos do que a maioria dos modelos estatísticos e econométricos, geralmente utilizados nas previsões do mercado (CUNHA et al., 2010). Porém, a principal dificuldade na utilização das redes neurais ocorre por serem modelos avaliados com pouca explicação das variáveis independentes em relação ao processo de predição, ou seja, “caixas pretas”.

Existe uma grande variedade de aplicações das redes neurais para a previsão do mercado financeiro. Entretanto, em consequência do comportamento aleatório e altamente complexo do mercado acionário, as redes neurais ainda são consideradas modelos matemáticos mais adequados para a modelagem do mercado de ações. Na grande maioria dos casos, as redes neurais são utilizadas para previsão dos preços e volumes no mercado acionário em função do mapeamento não linear entre as entradas e as saídas da rede (BRESSAN, 2004; OLIVEIRA et al., 2010). Acrescenta-se ainda que algumas redes utilizam como informações de entrada resultados da análise fundamentalista (GRAHAM e DODD, 2009), outras baseiam-se em séries temporais como o preço da ação e volume das negociações (GALÃO et al., 2011), contudo, o uso de sistemas de previsão como as redes neurais empregando variáveis da análise técnica como variáveis de entrada não tem sido estudado.

Então, com base nas transferências de incertezas e no controle de risco nas grandes volatilidades nos períodos de crises econômicas causadas pela grande imprevisibilidade do mercado tomado de decisões emocionais, questiona-se: Um método de previsão como as redes neurais, baseado em uma análise de mercado, pode ser utilizado em período de crises econômicas? E qual modelo de redes neurais tem melhor previsão em relação ao mercado acionário em épocas de crise financeira?

Com o intuito de auxiliar o investidor na tomada de decisão diante de uma crise econômica e visando contemplar o objeto deste estudo, descrito no parágrafo anterior, objetivou-se desenvolver duas redes neurais com aprendizados diferenciados sem o problema

da “caixa preta” dos métodos na linguagem JAVA no programa *Netbeans* e na linguagem C no programa Dev C++. Estas duas redes neurais serão desenvolvidas, treinadas e validadas para antever os movimentos do mercado quando este apresentar oscilações provenientes de uma crise econômica.

Dessa forma, o presente trabalho descreve dois sistemas que utilizaram dados em série temporais (PANDORFI, 2011) com o intuito de evidenciar qual modelo de rede neural, baseado em uma análise de mercado, pode ser utilizado em período de crises econômicas. Mais especificadamente, as saídas das redes neurais desenvolvidas deverão prever ao usuário oscilações provenientes de uma crise econômica, a partir do retorno do volume do Ibovespa e de algumas informações extraídas da análise técnica. Para validar o estudo, foi desenvolvida uma comparação entre os dois modelos de redes neurais demonstrando qual tem melhor previsão em relação ao mercado acionário em épocas de crise financeira, ou seja, foi comparado o poder de explicação dos dois métodos em relação à previsão de uma série temporal em um período de crise econômica.

Entre as redes neurais desenvolvidas, a primeira foi uma *Perceptron* de múltiplas camadas (MLP), treinada com o algoritmo de retropropagação de erro, com o aprendizado supervisionado (HAYKIN, 2001; SCHLKOFF, 2001; NASCIMENTO Jr. e YONEYANA, 2004; LUDWIG Jr. e MONTGOMERY, 2007; SOBREIRO et al., 2009; GALÃO et al., 2011). A segunda rede neural desenvolvida será uma rede **Kohonen**, treinada com o aprendizado não supervisionado e o algoritmo de retropropagação de erro (VALENÇA, 2009; MILIONI et al., 2007).

Como referido anteriormente, para validação das duas redes neurais, foram realizadas baterias de testes na predição do mercado acionário, nos anos de 2002 até 2011. Também foram realizados testes na comparação da predição das duas redes neurais já citados anteriormente, com o critério do menor Erro Médio Quadrático e o Coeficiente de Explicação (R^2) ajustado mais próximo de 1.

1.1 Delimitando o problema de pesquisa

Não há dúvidas de que uma das premissas básicas na dubiedade da decisão do ser humano minimize e/ou satisfaça seus anseios e necessidades. Entretanto, esta decisão nem sempre é tomada conscientemente, baseada em resultados de um método elaborado de forma racional. Em consequência disso, para uma boa tomada de decisão de um investidor do mercado acionário, devem-se levar em consideração, todos os elementos e variáveis

envolvidas, que permitam uma melhor gestão dos limites de riscos financeiros aceitáveis. Contudo, atualmente, esta abordagem da utilização de todos os fatores internos e externos de um projeto financeiro, como variável de entrada, seja pela impossibilidade física, ou pelo custo do processamento dos dados, torna-se inviável.

Porém, no gerenciamento de risco de uma carteira de ações, considera-se, que a utilização de métodos de previsão com variáveis provenientes do mercado acionário permitam um melhor controle, tanto quanto possível, dos fatores que interferem na tomada de decisão. Pois, geralmente, estes modelos automáticos de tomada de decisão, são baseados em critérios exatos, valores e/ou resultados lógicos que abrangem a maioria dos fatores impactantes. No entanto, a utilização de um método de previsão que utilize o retorno do volume semanal do Ibovespa para maximizar o retorno nem sempre é uma tarefa simples, pois alguns investidores *traders* acabam se baseando em métodos de previsão não muito convencionais e de soluções duvidosas (HISSA, 2008).

Como mencionado anteriormente, o foco principal deste trabalho é desenvolver, aplicar e validar dois tipos de redes neurais artificiais com os algoritmos de aprendizado diferentes que utilizem o retorno do volume semanal do Ibovespa como entrada de dados. A partir da abordagem utilizada, o trabalho busca identificar o melhor modelo, para a redução do risco na tomada da decisão do investidor *trader* (pessoa física) mediante oscilações provenientes de uma crise econômica no Ibovespa. Nesse contexto, surge o problema de pesquisa: Qual modelo de redes neurais desenvolvido, baseado em uma análise de mercado como variável de entrada, pode ser utilizado em período de crises econômicas? E que modelo de redes neurais tem melhor previsão em relação ao mercado acionário em épocas de crise financeira?

1.2 Objetivos

Para contemplar o objeto deste estudo, objetivou-se elaborar três grandes objetivos principais. Em primeiro lugar, desenvolver dois métodos heurísticos que utilizem ferramentas e estratégias de análise técnica com aprendizados diferenciados. Em seguida, eliminar o problema da “caixa preta” dos métodos desenvolvendo o algoritmo backpropagation de Garson (1991) modificado por Goh (1995) na linguagem JAVA no programa *Netbeans* para a primeira rede neural e na linguagem C no programa Dev C++ para a segunda rede neural. Em terceiro lugar, desenvolver, treinar e validar as duas redes neurais para antever os movimentos do mercado quando este apresentar oscilações provenientes de uma crise econômica.

Para desenvolver os três objetivos principais citado anteriormente, foi necessário delinear determinados objetivos específicos, sendo eles:

a) Desenvolver dois tipos de Redes Neurais Artificiais: Rede Neurais *Perceptron* de Múltiplas Camadas na linguagem JAVA no programa *Netbeans* e Redes Neurais **Kohonen** na linguagem C no programa Dev C++. Foi utilizada duas linguagens para o desenvolvimento das redes neurais, pois, tanto o desenvolvimento das duas redes neurais, quanto o domínio das linguagens foram em momentos distintos com um grande intervalos de tempo;

b) Identificar os pontos de possíveis crises econômicas na composição da saída da rede neural artificial, utilizado um *setup* com base na Teoria das Ondas de Elliott, contendo as nove regras dos princípios básicos de padrões e razões das ondas.

c) Modelar a série temporal do retorno do volume semanal do Ibovespa, levando-se em consideração a heterocedasticidade e o desenvolvimento dos dados ao longo do tempo em torno de valor constante, utilizando uma equação para estimar $Y_z^j(x_n)$ (variáveis independentes);

d) Na utilização do modelo de Garson (1991) modificado por Goh (1995), busca-se desenvolver as redes neurais a fim de obter uma melhor confiabilidade nos resultados obtidos. Este modelo permite ao algoritmo de correção do erro, o *backpropagation*, eliminar o problema da “caixa preta” das redes neurais, utilizando a influência total das variáveis dependentes nas variáveis independentes;

e) Comparar as previsões das duas redes neurais desenvolvidas e descobrir qual rede neural se adapta ao tipo de série temporal utilizada por meio dos testes da etapa de validação das redes neurais, que serão feitos ao término da verificação.

1.3 Justificativa

Quando investidores decidem investir no mercado de renda variável, como no mercado de ações, buscam um método que lhes dê mais segurança na tomada de decisão para controle do risco. Na prática, não há como saber quais ativos tornar-se-ão um investimento lucrativo, pois geralmente os sistemas utilizam métodos sem se abster de seus problemas originais.

No mercado acionário, os sistemas de previsão que utilizam um método heurístico como as redes neurais procuram auxiliar o investidor na tomada de decisões. Entretanto, estes métodos são compostos de vários problemas em relação à sua pouca explicação do resultado

das saídas com as variáveis de entrada. Por isso, geralmente são sistemas que não oferecem um resultado fiel em relação à previsão do mercado.

Porém, apesar da utilização de redes neurais para a tomada de decisão sem heterocedasticidade nos dados, sem o problema da “caixa preta” dos métodos e utilizando uma série com raiz unitária, ainda assim não se tem garantia de qual modelo terá melhor desempenho em relação à previsão do mercado em períodos de crises econômicas.

Em vista desta dificuldade, levando-se em consideração as afirmações anteriores, este trabalho colabora com um estudo inovador de finanças e um estudo contemporâneo de inteligência artificial de quatro principais maneiras. Em primeiro lugar, com a identificação de prováveis crises econômicas utilizando o retorno do volume semanal do Ibovespa a partir de um método altamente complexo como as redes neurais que possa ser utilizado por pesquisadores e investidores das áreas econômicas. Em segundo lugar, corroborando que, quanto mais complexo for o modelo e o método de treinamento da rede neural, melhor será seu poder de explicação. Ou seja, quanto mais robusto e complicado o modelo heurístico for, melhor desempenho terá em períodos de crises financeiras em comparação com o método de treinamento menos abstruso. Em terceiro lugar, dirá se a utilização da análise técnica no treinamento de um método heurístico pode auxiliar na sua previsão. E, em quarto lugar, indicará qual modelo de redes neurais tem melhor desempenho em relação à tomada de decisão no mercado financeiro, em períodos de crises econômicas.

1.5 Etapas do Trabalho

Este trabalho está estruturado em quatro capítulos, além desta breve introdução. O capítulo 2 apresenta a fundamentação teórica e alusões a respeito do Ibovespa e da natureza da crise econômica de 2007/2008. Também serão apresentados, neste capítulo, conceitos e citações sobre o desenvolvimento da rede neurais *Perceptron* de múltiplas camadas e redes neurais **Kohonen**, assim como seus problemas. Nos subcapítulos seguintes, será descrita a teoria em que as redes neurais se basearão para operar no mercado de capitais, assim como uma breve história dessa teoria. No capítulo 3, será apresentado os métodos utilizados para a realização do estudo, relatando como as redes neurais serão treinadas e validadas. O capítulo 4 oferece os resultados da modelagem dos métodos e os resultados da comparação das duas redes neurais, relatando qual método tem melhor previsão do mercado acionário em épocas de crises econômicas. O capítulo 5 relata as conclusões encontradas a respeito do objeto deste estudo, assim como aponta algumas alternativas, trabalhos futuros e limitações do estudo.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta um referencial teórico para o entendimento da natureza do Ibovespa e da crise financeira de 2007/2008, expondo-se uma breve história do Ibovespa e uma síntese a respeito das características e origens da crise financeira. Também são apresentados os conceitos e o desenvolvimento de duas redes neurais que serão utilizadas para a previsão do Ibovespa. Na parte final, é apresentado um levantamento dos problemas na utilização destes métodos e a resolução deles, bem como um breve conceito da teoria das ondas de Elliott utilizada no treinamento das redes neurais.

2.1 Ibovespa

Segundo Mara Luquet (2000), o principal índice do mercado acionário brasileiro é o Ibovespa. O fato fundamental que demonstra a sua importância é a fidelidade com que o índice retrata o comportamento dos principais papéis e *commodities* negociados na BM&FBOVESPA, garantindo uma tradição de integridade da série histórica, que não sofreu modificações metodológicas desde sua implantação, em 1968.

Então, este índice retrata o valor atual, em moeda corrente, de uma carteira teórica de ações que contém mais de 80% do número de negócios e do volume financeiro verificados no mercado à vista, representando aproximadamente 70% do somatório da capitalização de todas as empresas, com a finalidade básica de servir como indicador médio do comportamento do mercado (CAVALCANTI et al., 2005; BM&FBOVESPA, 2012).

Essa carteira teórica que forma o Ibovespa é composta pelas ações que dispõem cumulativamente de três critérios em relação a um período de doze meses anteriores à formação da carteira: estar incluídas em uma relação de ações cujos índices de negócios somados representem 80% do valor acumulado de todos os índices individuais; participar, em relação ao volume, superior a 0,1% do total; e terem sido negociadas em mais de 80% do total de pregões do período (BM&FBOVESPA, 2012).

2.2 Crises no Mercado Acionário

No tema estudado, o conceito de bolha pode ter muitos significados. Considerando os argumentos de não estabilidade e de equilíbrio, em finanças este conceito está relacionado a crises econômicas, ou seja, quando os valores presentes dos ativos não têm o mesmo fluxo de benefícios líquidos esperado pelos seus detentores (MARTIN et al., 2004). Em outras

palavras, as bolhas são relacionadas a crises econômicas pelo simples fato de serem fenômenos que ocorrem quando um ativo qualquer não concede os mesmos benefícios almejados depois da compra.

Do ponto de vista teórico, as bolhas são tidas como improváveis, mas a teoria convencional (neoclássica) não as nega e afirma que as mesmas só ocorrem em circunstâncias muito específicas (OREIRO, 2001). Porém, na prática, o conceito de bolha também está relacionado a mudanças e colapso dos mercados, que sofrem forte valorização ou desvalorização. Estas mudanças consideradas como movimentos imprevisíveis acontecem de tal forma que não se possa refletir sobre o valor justo dos ativos, evidenciando empiricamente que a teoria da eficiência do mercado oferece espaços para críticas (MARTIN et al., 2004).

Com base nos conceitos anteriores, as bolhas são consideradas evolutivas e decorrem de um processo endógeno de mudança (OREIRO, 2001). Desta forma, ponderando o processo endógeno de transformação, a primeira bolha da história foi a explosão do período cambriano, que ocorreu por volta de 530 milhões de anos. Para os estudiosos das áreas evolutivas, não houve precedentes, aconteceu em um instante e nos 70 ou 80 milhões de anos seguintes, a evolução se acelerou tanto que a raça humana surgiu (STONE et al., 2011). Nesse sentido, as bolhas são escalonadas e sistêmicas como as interligações dos mercados mundiais que ocorreram a partir da evolução da globalização dos mercados.

As interligações dos mercados mundiais se devem a vários fatores, mas, como mencionado anteriormente, o principal destes foi a globalização que se iniciou no final da década de 80, proporcionando a diversificação de investimentos ao redor do mundo ao mesmo tempo (BARTRAM e BODNAR, 2009). Porém, mesmo apresentando fatores positivos, a globalização também acarretou alguns negativos. Em relação a uma crise econômica, não somente ocasiona incertezas ao seu respectivo mercado acionário, mas também aos mercados que estão intrinsecamente ligados a ele, conforme Barba (2011).

O grande conceito que envolve as crises econômicas é a multiplicidade de episódios nos quais instituições ou ativos financeiros se desvalorizam abruptamente (KINDLEBERGER e ALIBER, 2005). Estes fenômenos, denominados crises econômicas, são geralmente associados às corridas aos bancos, a bolhas financeiras e à quebra de mercados de ações (LAEVEN e VALENCIA, 2008).

A maior dificuldade para se identificar a data de emergência de uma crise econômica está relacionada aos vários acontecimentos políticos, macroeconômicos e microeconômicos ao redor do mundo. As crises econômicas são consideradas como fenômenos que, segundo Huyghebaert e Wang (2010), corroborando com Khalid e Rajaguru (2007 apud BARBA,

2011), são compostos por padrões dos movimentos negativos dos mercados de capitais. Estes padrões de movimentos mudam significativamente depois de grandes eventos econômicos imprevisíveis, impactando negativamente nos mercados de capitais ao redor do mundo (BARBA, 2011).

Quanto à ocorrência de uma crise econômica, a partir dos dados de Laeven e Valencia (2008), pode-se concluir que, em média, a cada ano incidiram três crises bancárias sistêmicas, cinco crises cambiais e um episódio de não pagamento de dívida soberana desde 1970 até os dias atuais. Mas foi a partir de 1900, mais especificadamente na Grande Depressão, na década de 1930, que a frequência das crises econômicas aumentou esporadicamente, deixando os investidores mais apreensivos a cada ano.

2.2.1 Crise da Inadimplência do *Subprime* 2008

Segundo Laeven e Valencia (2008), que apresentam um banco de dados sobre todas as crises bancárias, no período de 1970 a 2007, o agente causador da crise de 2007/2008 teve origem financeira. A estimativa de perdas da crise internacional de 2008, intitulada a Crise da Inadimplência do *Subprime*, atingiu seu auge em março de 2008, com a quebra do 5º maior banco de investimentos americano, o Bear Stearns (TORRES e BORÇA, 2008).

Até este período, a crise do *Subprime* já tinha atingido um *déficit* de 380 bilhões de dólares nos bancos europeus e americanos que participaram dos empréstimos imobiliários (TORRES, 2008). Somente para se ter uma ideia, segundo Soares (2009), o *déficit* de 955 bilhões da crise de 2007/2008 foi 297% maior que o Produto Interno Bruto anual (PIB) do Uruguai, em meados de 2007 (IPEADATA, 2011), configurando-se como uma das maiores crises econômicas já ocorridas nos últimos tempos.

Entretanto, para interpretar o processo de endividamento dos bancos europeus e americanos, deve-se considerar que a crise iniciou-se muito antes do seu auge. Segundo Torres (2008), a crise do *Subprime* começou com a transferência de diferentes contratos de hipotecas para um único fundo de investimento, em 2003. Cada contrato tinha um retorno em relação ao risco, ou seja, quanto maior o retorno, maior o risco. Os riscos dos contratos eram calculados a partir de modelos estatísticos validados pelas principais empresas de classificação de riscos do mundo, aparentemente seguros.

Dessa forma, as empresas de classificação de risco classificavam esses contratos em AAA, AA, A, BBB, BB, B e lixo tóxico (*toxic waste*), sendo os contratos AAA os menos arriscados e o mais arriscado, o lixo tóxico (TORRES, 2008). Mas, para a diluição de risco,

alguns contratos AAA eram somados a contratos lixo tóxico, transformando-os em AA ou A ou BBB e vendidos a fundos CDO (*collateralized debt obligations* – “títulos garantidos por dívidas”), junto com outros títulos da dívida, derivativos e até mesmo outros investimentos imobiliários, formando uma cascata.

Com uma política de classificação do crédito fraca em relação aos tomadores de empréstimos *Subprime*, os contratos de hipotecas acabaram tornando-se cada vez mais inadimplentes e assim todos os contratos foram sendo depreciados em cascata, até mesmo os contratos com os menores riscos (LAEVEN e VALENCIA, 2008). Então, pela falta de informações claras sobre o tamanho e a responsabilidade das perdas, os investidores pararam de renovar e começaram a resgatar suas aplicações em fundos imobiliários CDO's (BLANCHARD, 2008). Isso ocasionou um aumento dos custos de captação, inclusive nos mercados interbancários, transferindo os riscos dos mercados imobiliários americanos aos bancos europeus e americanos (TORRES, 2008).

Para Blanchard (2008), os dois grandes mecanismos do colapso dos empréstimos *Subprime* foram a venda de ativos para satisfazer as corridas de liquidez por parte dos investidores e a venda de ativos para restabelecer a divisão de capital. Em seu estudo, Blanchard concluiu que, se as políticas monetárias governamentais de curto prazo não mudarem, as oscilações no mercado acionário por meio de crises econômicas, riscos de deflação e armadilhas de liquidez com a finalidade de retirar o capital dos países emergentes continuarão ocorrendo como um ciclo.

Porém, nos dias atuais, existem métodos de previsão muito robustos e computacionalmente poderosos, a fim de prever as oscilações do mercado de renda variável com maior realidade possível, com o intuito de controlar o risco na tomada de decisão do investidor pessoa física. Um destes métodos heurísticos são as redes neurais, consideradas por muitos pesquisadores como Oliveira et al., (2010) e Campos, (2008) os melhores métodos de previsibilidade no mercado acionário em relação a modelos estatísticos e econométricos.

2.3 Breve Historia a Respeito das Redes Neurais Artificiais

A ideia inicial, no século XIX, com a revolução industrial, o advento da tecnologia e o avanço dos processos automáticos, era obter um dispositivo capaz de substituir o homem em uma ou várias atividades cotidianas com características não lineares (LUDWIG Jr. e MONTGOMERY, 2007). Mas foi o desenvolvimento de técnicas de simulação do processo de aprendizado, raciocínio e conhecimento que fez muito destes objetivos serem atingidos.

Em meio a estes se encontram as redes neurais artificiais, que proporcionaram, na década de 40, uma nova era de cognição e inteligência às máquinas (VALENÇA, 2009).

Este instrumento foi desenvolvido idealizando ampliar uma analogia entre o neurônio biológico e os circuitos elétricos, baseada na simulação de conexões sinápticas pelo uso de resistores¹ e amplificadores² (VALENÇA, 2009; LUDWIG Jr. e MONTGOMERY, 2007). Logo, as redes neurais são consideradas como uma das técnicas de inteligência artificial mais antigas em uso, pelo fato da sua grande capacidade computacional (LUDWIG Jr. e MONTGOMERY, 2007).

Uma rede neural artificial é composta por vários neurônios matemáticos que formam um modelo formal matemático para replicar o funcionamento e o processo de aprendizagem de uma cadeia de neurônios biológicos (JUDD, 1990). O desenvolvimento de uma rede neural artificial compreende diversos conceitos correspondentes a áreas distintas (LUDWIG Jr. e MONTGOMERY, 2007). Um dos grandes exemplos disso foi a contribuição da Neurofisiologia com os estudos das conexões entre neurônios biológicos e o momento da aquisição do conhecimento no cérebro humano. A Matemática forneceu a construção do neurônio matemático. E a Computação, no final do século XIX, mesclou estas técnicas, automatizando-as, criando a Neurocomputação, área que estuda as redes neurais artificiais como um meio de transferir o aprendizado humano para sistemas automatizados.

2.4 Neurônio Biológico

Alguns conceitos básicos referentes ao funcionamento do cérebro humano e seus componentes, os neurônios, são de fundamental importância para compreender a lógica de funcionamento das redes neurais (LUDWIG Jr. e MONTGOMERY, 2007). De maneira bastante simplificada, no neurônio biológico, há três componentes bastante significantes: os dendritos, os axônios e o corpo celular (VALENÇA, 2009).

Os dendritos têm a função de receber os estímulos transmitidos pelos outros neurônios. Entretanto, apesar do nome, no neurônio biológico, estes estímulos não são uma corrente elétrica, mas uma corrente de íons de sódio e potássio que caminham para os axônios contendo informações sobre um determinado estímulo corporal (LUDWIG Jr. e MONTGOMERY, 2007). Portanto, os axônios são responsáveis pelos estímulos recebidos e

¹ Resistores são dispositivos que transformam energia elétrica em energia térmica (WEISZFLOG, 2009).

² Amplificadores são dispositivos que ampliam uma energia menor para controlar uma energia maior (WEISZFLOG, 2009).

transmitidos para as células, e o processo de ligação de um axônio a um dendrito é composto por uma *sinapse* que se comunica por meio da substância serotonina. Já o corpo celular é responsável pela coleta e combinações de informações vindas de outros neurônios, e neste processo um sinal pode ser amplificado ou atenuado (NASCIMENTO Jr. e YONEYANA, 2004).

O ciclo de funcionamento de um neurônio biológico fundamentalmente é composto por três etapas: a primeira consiste em receber os estímulos enviados pelas *sinapses* dos sensores (papilas gustativas, pele etc.) seguindo pelo axônio, se o sinal for superior ao limiar até, o corpo celular (LUDWIG Jr. e MONTGOMERY, 2007); a segunda consiste em compilar e calcular o estímulo recebido no corpo celular; já a terceira etapa envolve repassar, pelos dendritos, o resultado calculado pelo corpo celular para o próximo neurônio (VALENÇA, 2009).

2.5 Neurônio Matemático

Percebendo o funcionamento e baseado nos principais componentes de um neurônio biológico, Warren McCulloch, neurofisiologista, e Walter Pitts, matemático, de forma simples, desenvolveram o neurônio matemático, representando-o por meio de uma regra de propagação e uma função de ativação (VALENÇA, 2009). Eles combinaram os conceitos trazidos pela Neurofisiologia, como o processo de aprendizagem de um neurônio biológico, e pela Matemática, como pesos e *thresholds*, para desenvolver o neurônio matemático (MACCULLOCH e PITTS, 1943).

No neurônio matemático, as *sinapses* fazem a parte dos dendritos do neurônio biológico. No modelo matemático, o papel dos axônios é desempenhado, pelos *bias* (LUDWIG Jr. e MONTGOMERY, 2007). No neurônio biológico, o corpo celular é responsável pelo processamento dos estímulos. Já no neurônio matemático esta função é responsabilidade da função de transferência, Equação (2.1) e da função de ativação, Equação (2.2) (FARIA et al., 2008; LUDWIG Jr. e MONTGOMERY, 2007). O processo de funcionamento de uma rede neural artificial, similarmente ao natural, consiste em receber um ou mais estímulos de entrada $x_1, x_2, x_3, \dots, x_n$, como sendo as variáveis de entrada x_n ($n=1, \dots, n$) e devolver um único sinal z de saída da rede para o neurônios da camada posterior. A regra de propagação é dada pela entrada líquida V_z :

$$V_z = \sum_{j=1}^z w_{ij} \cdot x_n. \quad (2.1)$$

Na nomenclatura utilizada, w_{ij} são os pesos sinápticos, de modo que a letra j representa o neurônio que recebe o sinal (saída); a letra i , o neurônio que emite o sinal (entrada); e a letra n concebe o número de entradas. A Figura 1 exibe esta representação do neurônio matemático, ilustrando os principais componentes de uma rede neural artificial presentes em Ludwig Jr. e Montgomery (2007), Nascimento Jr. e Yoneyana (2004).

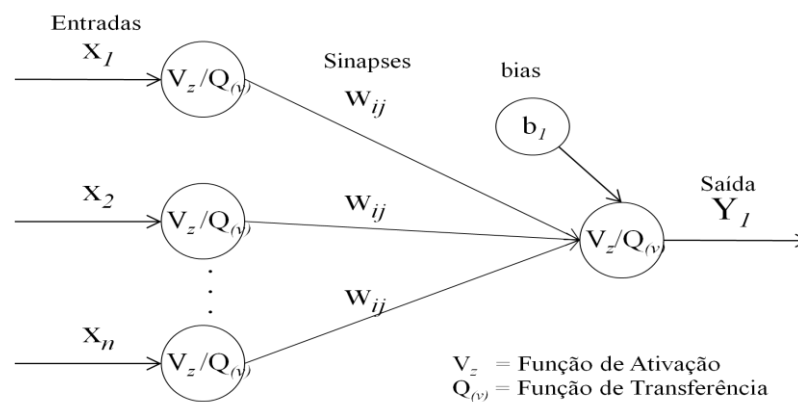


Figura 1 - Neurônio Matemático

Fonte: Ludwig Jr. e Montgomery (2007); Nascimento Jr. e Yoneyana (2004); Ceretta et al. (2010).

Na Figura 1, são definidas as entradas em x_n , as *sinapses* em w_{ij} , o *bias* em b_1 e em Y_1 a saída $f(V_z)$ de cada neurônio da camada de saída. O *bias* b_n , e as *sinapses* w_{ij} são modificados a cada interação, até chegar à saída desejada ou completar os ciclos escolhidos. Com a utilização do *bias*, aumentam os graus de liberdade da função de transferência, o que, consequentemente, aumenta a capacidade de aproximação da rede neural, sem saídas nulas (LUDWIG Jr. e MONTGOMERY, 2007). As funções de ativação e transferência são responsáveis pela passagem do estímulo e pelo cálculo feito dentro do neurônio.

2.5.1 Função de Ativação

A função de ativação (2.1) descrita em Ludwig Jr. e Montgomery (2007), é responsável por simular a primeira etapa dentro do neurônio, onde é feito o cálculo de ativação da estrutura por meio de uma função que está simulada pela letra V_z , Expressão 2.1.

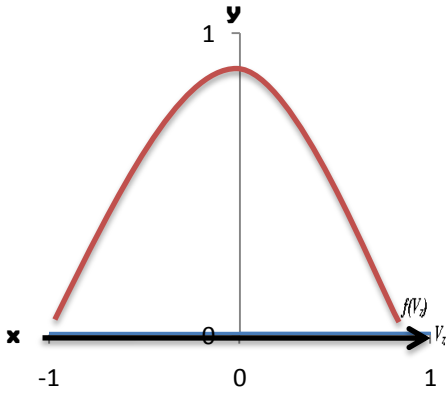
Esta estrutura é desenvolvida dentro do neurônio matemático para corresponder ao primeiro cálculo dos estímulos que chegam pelos dendritos de um neurônio biológico.

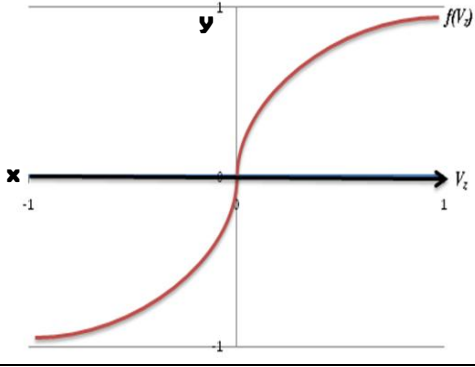
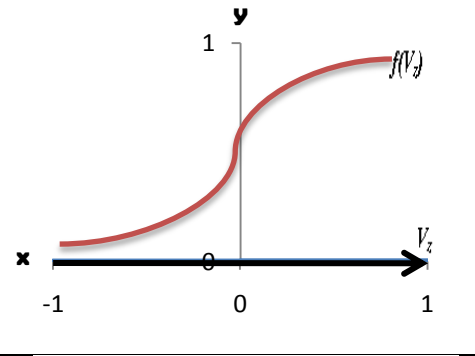
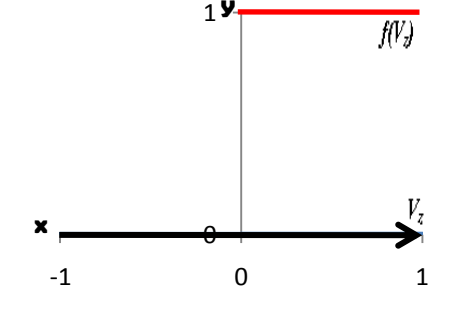
Na Expressão (2.1), o estímulo de todas as entradas recebidas pelo neurônio e os bias correspondentes são somados e multiplicados pelo resultado das camadas seguintes. Posteriormente, é passado o resultado para a segunda etapa dessa fase, a de transferência, que corresponde ao funcionamento do corpo celular por meio de uma regra de propagação (LUDWIG Jr. e MONTGOMERY, 2007).

2.5.2 Função de Transferência Regra de Propagação

Como mencionado anteriormente, o corpo celular biológico tem a função de coletar, calcular e repassar o estímulo aos próximos neurônios. Em redes neurais com mais de uma camada, a regra de propagação é responsável pela função do corpo celular biológico (VALENÇA, 2009). Nesta regra de propagação, dá-se uma função de transferência não linear, que tem, dentre outras, o escopo de evitar que os valores de saída afastem-se do intervalo de valores máximos e mínimos determinados.

Portanto, a responsabilidade da transferência do estímulo recebido e emitido entre os neurônios matemáticos é auferida a quatro diferentes funções mais destacadas nas redes neurais, cada uma delas com suas características distintas. Mas, dependendo das características dos dados, as funções de ativação que melhor se adaptam na prática para uma distribuição não binomial são a função sigmóide logística e a função tangente hiperbólica (GOMES e LUDERMIR, 2008). O Quadro 1, abaixo, demonstra as quatro funções mais utilizadas com os seus respectivos gráficos (LUDWIG Jr. e MONTGOMERY, 2007).

Função	Gráfico
<p style="text-align: center;">Gaussiana (1) $Q(V_z) = f(V_z) = e^{-V_z^2}$</p>	

<p>Tangente Hiperbólica (2)</p> $Q(V_z) = f(V_z) = \frac{e^{(V_z)} - e^{-(V_z)}}{e^{(V_z)} + e^{-(V_z)}}$	
<p>Função</p>	<p>Gráfico</p>
<p>Sigmóide (3)</p> $Q(V_z) = f(V_z) = \frac{1}{1 + e^{-(V_z)}}$	
<p>Degraus (4)</p> $Q(V_z) = f(V_z) = \begin{cases} 1, & V_z \geq 0 \\ 0, & V_z < 0 \end{cases}$	

Quadro 1 - Funções de Transferência

Fonte: Ludwig Jr. e Montgomery (2007); Valença (2009)

Se a função de transferência fosse linear, na busca de um erro mínimo quadrático, a derivada seria uma constante sem nenhuma informação sobre o valor pelos pesos *sinápticos*, mas, ao se utilizar uma técnica do gradiente como a função sigmóide, elimina-se o problema do ajuste dos pesos *sinápticos* com série constante (LUDWIG Jr. e MONTGOMERY, 2007). Por isso, a função de transferência utilizada nas redes neurais desenvolvidas está simulada na Expressão (2.2), a seguir. Nesse caso, a Expressão utilizada é a sigmóide. Nesta função, $Q(V_z)$ é o valor do resultado da função de transferência; V_z é o valor do resultado da função de ativação; e é a base natural de logaritmos.

$$Q(V_z) = \frac{1}{(1 + e^{-(V_z)})}. \quad (2.2)$$

Considerando que os valores de entrada são $x_0 =$ constante que segura a função de normalização, ou seja, função de transferência; $x_1 = -1$ e $x_2 = 1$ os pesos sinápticos $w_{10} = -1$; $w_{11} = 3$ e $w_{12} = 4,5$, a saída da função de transferência sigmóide da rede será:

$$V_z = w_{10}.x_0 + w_{11}.x_1 + w_{12}.x_2;$$

$$V_z = -1.(1) + 3.(-1) + 4,5.(1) = 0,5.$$

a) Cálculo da saída para a função sigmóide.
onde $V_z = 0,5$.

$$Y = Q(V_z) = \frac{1}{(1 + e^{-(0,5)})};$$

$$Y = Q(V_z) = 0,62.$$

Em caso de aprendizado supervisionado, se a saída não for a almejada, compara-se com a saída desejada $d_j(n)$, obtendo-se o valor da Expressão de erro $e_j(n)$, descrita mais detalhadamente no subcapítulo 2.3.4.1. Se o aprendizado for não supervisionado, para a escolha do neurônio vencedor, calcula-se a distância euclidiana d_{xw} entre os pesos *sinápticos* e as entradas, descrita no subcapítulo 2.3.5.1, onde se verão mais minuciosamente as relações (VALENÇA, 2009).

2.5.3 Redes Neurais *Perceptron*

Segundo Yonenaga e Figueiredo (1999), em 1958, Frank Rosenblatt construiu um dos primeiros tipos de redes neurais, a rede neural artificial *Perceptron*. No seu trabalho intitulado “*The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*”, Rosenblatt, desenvolveu uma simulação computacional que distribuía uma única classe de entrada em um conjunto de classes pré-estabelecidas. Com este trabalho, corroborou que o reconhecimento de padrões é uma das características do *Perceptron*.

Esta rede desenvolvida por Frank Rosenblatt atualiza as sinapses no treinamento com um método de ajuste supervisionado, quando o supervisor ajusta os parâmetros modificáveis da rede neural com base no valor do erro (YONENAGA e FIGUEIREDO, 1999). A menor

representação da classe *Perceptron* como ajuste supervisionado é um neurônio na camada de entrada, ligado a um neurônio na camada de saída.

Mas, logo após a publicação do trabalho de Frank Rosenblatt, em 1958, Marvin Minsky e Seymour Papert descobriram e provaram que as redes neurais de uma única camada não são tão eficientes assim (LUDWIG Jr. e MONTGOMERY, 2007). Com a publicação do trabalho de Minsky e Papert, desestimularam-se muitos pesquisadores da época na área de redes neurais (NASCIMENTO Jr. e YONEYANA, 2004). No trabalho intitulado “*Perceptrons*”, os autores corroboraram que a utilização da função degraus pela *Perceptron* de Frank Rosenblatt fornece somente um resultado com apenas um sinal circunspecto (sim ou não) e, por consequência, acarreta o problema da não resolução da operação do (OU) exclusivo (XOR) como função linearmente separável (VALENÇA, 2009). Ou seja, a rede neural *Perceptron* de Frank Rosenblatt não consegue fornecer saídas para alguns problemas com dissoluções não binárias.

Para exemplificar o problema, apresenta-se a função degraus, a seguir, com dois exemplos de funções lógicas solucionáveis pelas redes *Perceptron* e um exemplo de função lógica não solucionável que corresponde ao problema da função não linearmente separável do (OU) exclusivo (XOR). A primeira função lógica solucionável pela *Perceptron* é a função lógica “E” apresentada na Tabela 1, a seguir, onde se multiplicam as entradas para obter os resultados da função. Ou seja, x_1 multiplicado por x_2 , igual a Y .

Tabela 1- Função lógica “E”

Entradas		Saída		
x_1	x_2		=	Y
0	.	0	=	0
0	.	1	=	0
1	.	0	=	0
1	.	1	=	1

Fonte: Valença (2009)

O gráfico desta função lógica linearmente separável, Tabela 1, pode ser visualizado na Figura 2, onde há uma reta capaz de separar as classes “1” e “0”. Quando isso é possível, diz-se que a função é linearmente separável (VALENÇA, 2009). Ou seja, os quatro pontos traçados no gráfico correspondentes às duas classes podem ser separados por somente uma reta.

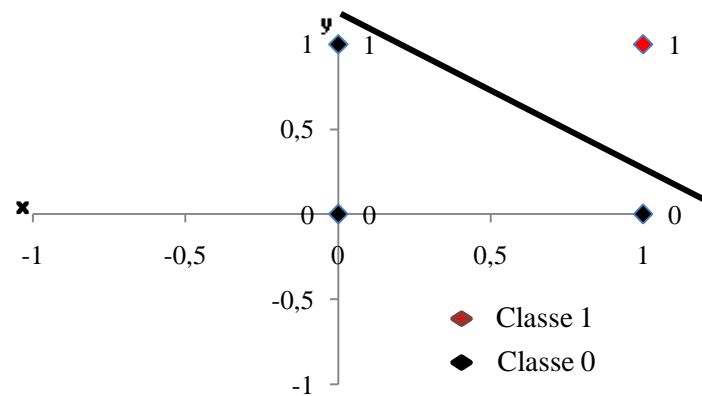


Figura 2 - Gráfico da função lógica “E”

Fonte: Organizada pelos Autores

A segunda função lógica solucionável pela *Perceptron* é a função lógica “OU” apresentada na Tabela 2. Esta tabela contém as entradas e os seus respectivos resultados na utilização da função lógica “OU”, que corresponde ao somatório das entradas. Ou seja, x_1 , mais o operador aritmético $+$, x_2 , igual a Y .

Tabela 2 - Função lógica “OU”

Entradas			Saída	
x_1		x_2	=	Y
0	+	0	=	0
0	+	1	=	1
1	+	0	=	1
1	+	1	=	1

Fonte: Valença (2009)

O gráfico desta função linearmente separável pode ser visualizado na Figura 3, a seguir, onde existe uma reta capaz de separar as classes “1” e “0”. Quando isso é possível, diz-se que a função é linearmente separável (VALENÇA, 2009). Ou seja, os quatro pontos traçados no gráfico correspondentes às duas classes podem ser separados por somente uma reta.

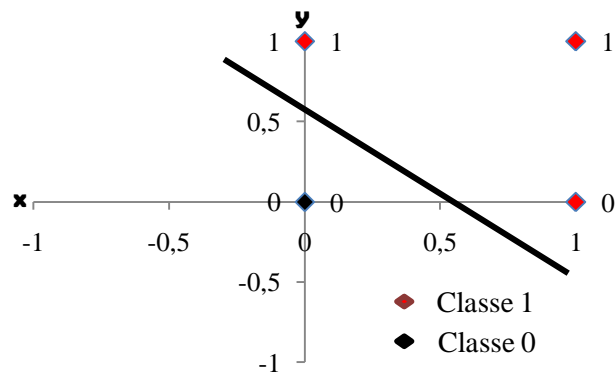


Figura 3 - Gráfico da função lógica “OU”

Fonte: Organizada pelos autores

A terceira função lógica (OU) exclusivo “XOR”, apresentada na Tabela 3, abaixo, obedece a problemas não solucionáveis pelas redes neurais *Perceptron*. Esta tabela contém as entradas e os seus respectivos resultados na utilização da função lógica “OU” exclusivo “XOR”, que corresponde à negação do somatório das entradas. Ou seja, x_1 , mais a negação do operador aritmético $+$, x_2 , igual a Y .

Tabela 3- Função lógica OU exclusivo “XOR”

Entradas		Saída		
x_1		x_2	=	Y
0	\oplus	0	=	1
0	\oplus	1	=	0
1	\oplus	0	=	0
1	\oplus	1	=	1

Fonte: Valença (2009)

O gráfico desta função linearmente não separável pode ser visualizado na Figura 4, onde se pode verificar que não existe uma única reta que seja capaz de separar as classes “1” e “0”, assim, a função é dita linearmente não separável (VALENÇA, 2009). Ou seja, os quatro pontos traçados no gráfico correspondentes às duas classes classificadas 1 e 0 não podem ser separados por somente uma reta, considerando-se o problema como não resoluto por uma rede *Perceptron*.

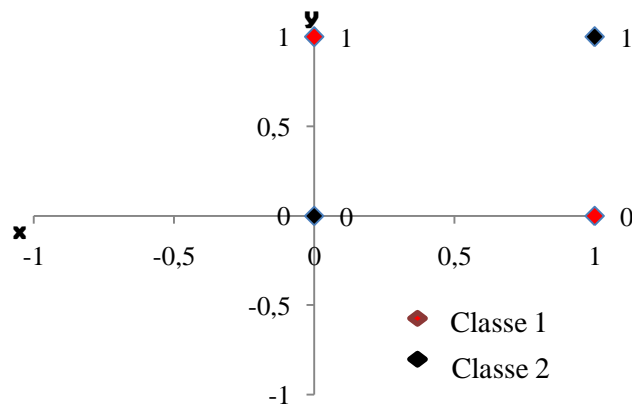


Figura 4 - Gráfico da função lógica OU exclusivo “XOR”

Fonte: Organizada pelos autores

Nas Tabelas 1 e 2, verifica-se que as redes *Perceptrons* são capazes de classificar correta e linearmente separáveis os valores de entrada inicialmente adotados x_1 e x_2 e o conjunto de pesos encontrados w_{ij} . Portanto, se existe algum conjunto de soluções possíveis, o *Perceptron*, durante o treinamento, irá procurar uma destas soluções, mas se não houver resultados plausíveis, como demonstrado no exemplo do problema da função não linearmente separável do OU exclusivo (XOR), o *Perceptrons* não achará nenhum resultado.

2.5.4 *Perceptron* Múltiplas Camadas (MLP)

Como visto no subcapítulo anterior, o problema do (OU) exclusivo (XOR) na utilização da *Perceptron* simples inviabilizou os estudos e os processos de inteligência artificial com o passar do tempo (LUDWIG Jr. e MONTGOMERY, 2007). Mas, logo após a descoberta de Marvin Minsky e Seymour Papert em relação ao problema das redes *Perceptron*, John Holpfield, em 1982, estabeleceu uma descrição das propriedades associativas das redes neurais de múltiplas camadas, levando à estruturação de uma nova estratégia (VALENÇA, 2009).

Nesta nova estratégia, propôs uma rede neural mais robusta, com algumas camadas entre a de saída e a de entrada, denominadas camadas ocultas. Com a criação desta nova técnica, eliminou-se completamente o antigo problema do (OU) exclusivo (XOR), induzindo, na época, a retomada do interesse nos estudos das áreas que compreendem esta ferramenta matemática (NASCIMENTO Jr. e YONEYANA. 2004). Este formalismo com camadas de neurônios internas denominou-se de *Multilayer Perceptron*, estudado nos próximos subcapítulos.

2.5.4.1 Definição de uma MLP

Como mencionado anteriormente, essa ferramenta, denominada *Multilayer Perceptron*, constitui-se de camadas de entrada sensoriais com tantos neurônios quantos forem os sinais de entrada, uma ou várias camadas ocultas e uma camada de saída, com o número de neurônios igual ao número de pesos de saída (LUDWIG Jr. e MONTGOMERY, 2007; VALENÇA, 2009; GONÇALVES, 2005; NASCIMENTO Jr. e YONEYANA, 2004). Com esta adição de pelo menos uma camada escondida, as redes MLP proporcionaram uma grande vantagem na generalização sobre as *Perceptrons* simples.

Um dos grandes benefícios na utilização de uma rede neural MLP diz respeito à extrema precisão em suas respostas, que permite resolver problemas reais (funções não linearmente separáveis). Mas a grande vantagem de utilizá-las é o conceito de que uma rede neural imita o cérebro humano, isto é, transcreve os atos de aprendizagem com erros e acertos em suas tentativas (ANITHA et al., 2012). Porém as redes neurais MLP, na prática, não conseguem fornecer soluções para problemas muito complexos sozinhas. Então, o algoritmo de retropropagação do erro com aprendizagem supervisionada tem o papel de ajudá-las nessa tarefa, decompondo os problemas complexos em um número de tarefas relativamente simples (LUDWIG Jr. e MONTGOMERY, 2007; SCHALKOFF, 2001).

2.5.4.2 Retropropagação do Erro com Aprendizagem Supervisionada

Há vários algoritmos que produzem o aprendizado para as redes neurais artificiais (LUDWIG Jr. e MONTGOMERY, 2007). O algoritmo *backpropagation* de correção do erro com aprendizagem supervisionada é um deles. Este algoritmo é utilizado para controlar o erro e a taxa de aprendizado em redes neurais que necessitam de um supervisor como as *Perceptrons*.

Este treinamento supervisionado em relação à sua aplicação e ao seu desenvolvimento pode ser classificado de duas formas totalmente distintas: o método de treinamento com supervisor forte e o método de treinamento com supervisor fraco. Então, quando a supervisão for forte, ele fornece a saída desejada para a rede neural conforme for o seu propósito, logo, aprendizagem supervisionada. Quando o supervisor não ajusta o erro do algoritmo pela saída desejada, ele é chamado de supervisão fraca, portanto, o algoritmo de aprendizagem é não supervisionado.

Como referido anteriormente, o algoritmo utilizado pelas redes neurais *Perceptron* é o de retropropagação de erro com aprendizagem supervisionada, pois, neste tipo de algoritmo, o supervisor fornece a saída desejada e a controla a partir do erro da rede. No entanto, segundo Valença (2009), para representar o desenvolvimento do funcionamento deste algoritmo, depende-se de quatro passos.

O primeiro passo inicia-se com o vetor de valores de entrada $x(n)=[x_1, x_2, \dots, x_n]$, na primeira camada de neurônios. No segundo passo, são auferidos valores aleatórios para as *sinapses* w_{ij} , e *bias* b_n , da rede a fim de realizar os primeiros cálculos de ativação V_z , e de transferência $Q(V_z)$ de cada neurônio, repassando o resultado à entrada do neurônio na camada seguinte, prosseguindo até chegar à última camada de neurônios. No terceiro passo, na última camada de neurônios, é calculada a diferenciação $e_f(n)$, entre a saída desejada $d_f(n)$, e a saída calculada pela rede neural $Y_z^j(n)$. Como mencionado anteriormente, esse processo de comparação é chamado de aprendizagem supervisionada, com o supervisor forte, pois, além de manipular a saída desejada $d_f(n)$, da rede, o supervisor fornece um ajuste dos pesos com base no erro $e_f(n)$. O quarto passo somente será desenvolvido se o resultado do erro não for satisfatório, pois é neste passo que ocorrerá a correção das *sinapses* w_{ij} , e *bias* b_n , dos neurônios.

Portanto, simplificados estes passos, pode-se chegar a basicamente dois instantes O primeiro passo (Fase Forward) refere-se à progressão do sinal do primeiro neurônio até a última camada de neurônios. Já o segundo passo (Fase Backward) refere-se à propagação do erro da última camada da rede até o primeiro neurônio da camada de entrada, corrigindo os pesos *sinápticos* w_{ij} , e *bias* b_n (VALENÇA, 2009), ou seja, do último neurônio da última camada para o primeiro neurônio da primeira camada (LUDWIG Jr. e MONTGOMERY, 2007). Esta correção, chamada de retropropagação do erro, ocorrerá se o cálculo do erro do neurônio de saída e a taxa de aprendizado não forem satisfatórios.

2.5.4.2.1 Propagação progressiva do sinal (Fase Forward)

Como se viu no subcapítulo anterior, o primeiro passo do algoritmo de retropropagação do erro com aprendizagem supervisionada consiste em realizar a propagação dos pesos da camada de entrada para a camada de saída (VALENÇA, 2009), calculando-se as entradas da Fase Forward. Supondo-se uma rede neural que utilize a função de transferência sigmóide, Expressão (2.3), com três camadas $Y_1^j = Q(V_z^j)$, $Y_2^j = Q(V_z^j)$, $Y_3^j = Q(V_z^j)$, com três neurônios na camada escondida, incluindo o neurônio da constante e um neurônio na camada de saída,

onde j se refere ao número do neurônio na camada, de tal forma que w_{ij} ($j=1, 2, 3$); e z refere-se aos números das camadas da rede; de tal forma que V_z^j ($z=1, \dots, 3$), apresenta-se:

$$Y_z^j = Q(V_z^j) = \frac{1}{\left(1 + e^{-\left(\sum_{j=1}^z w_{ij} \cdot x_n\right)}\right)}. \quad (2.3)$$

Para o exemplo, ter-se-ia que calcular sete funções de transferência $Q(V_z^j)$, uma para cada neurônio da rede neural e sete funções de ativação V_z^j correspondentes aos mesmos sete neurônios, completando a primeira fase do algoritmo (Fase Forward) (NASCIMENTO Jr. e YONEYANA. 2004). Então, têm-se três saídas correspondentes aos três primeiros neurônios, que são:

$$Y_1^1 = Q(V_1^1), Y_1^2 = Q(V_1^2), Y_1^3 = Q(V_1^3) = 1.$$

Para os três neurônios da camada escondida, que são:

$$Y_2^1 = Q(V_2^1), Y_2^2 = Q(V_2^2), Y_2^3 = Q(V_2^3) = 1.$$

E para o último neurônio na camada de saída, que é:

$$Y_3^1 = Q(V_3^1).$$

Com o desenvolvimento destas 14 funções, tem-se a fase de propagação progressiva do sinal completa, ou seja, a propagação dos pesos de entrada difundindo-se até o último neurônio da camada de saída. Entretanto, como mencionado anteriormente, para que a aquisição do conhecimento seja empregada às *sinapses*, é necessária a utilização da fase de reajuste dos pesos descrita no próximo subcapítulo.

2.5.4.2.2 Reajuste dos Pesos (Fase Backward)

Na segunda fase do algoritmo, ocorrerá a retropropagação do erro com aprendizagem supervisionada de modo que o processo de retropropagação da Fase Backward incida em paralelo com o reajuste das *sinapses* w_{ij} , e *bias* b_n , da rede neural MLP (VALENÇA, 2009). O

grande propósito da utilização deste processo de retropropagação consiste em ajustar a previsão de acordo com o erro disponibilizado pelo usuário e o erro entre a saída desejada e a saída calculada pela rede neural. Portanto, o ajuste dos pesos da rede neural relaciona-se com o erro entre os dois parâmetros, saída desejada $d_j(n)$ e a saída calculada $Y_j(n)$. Se não forem aceitáveis, novos pesos para as *sinapses* w_{ij} , e *bias* b_n , do neurônio serão gerados de modo que este erro não possa ser inferior a um erro escolhido pelo usuário ou ocasionará uma parada forçada da rede neural, e novos valores terão que ser selecionado (LUDWIG Jr. e MONTGOMERY, 2007; GUIMARÃES, 2008).

Em uma rede neural de múltiplas camadas, utiliza-se o valor de erro para definir uma parada aceitável. A rede neural artificial procurará aproximar-se desse valor ou até mesmo igualá-lo em alguns casos. Logo, para medir o desempenho do treinamento, diminui-se o valor desejado do valor calculado do neurônio. A diferença indica o quanto a rede aprendeu com o treinamento supervisionado e, somente se seu erro ou desempenho for aceitável, será considerada treinada (VALENÇA, 2009; LUDWIG Jr. e MONTGOMERY, 2007). Mas a correção dos pesos incidirá do último neurônio da última camada até o primeiro neurônio da primeira camada da rede neural, levando-se em consideração o resultado obtido no cálculo do erro $e_j(\mathbf{n})$, Expressão (2.4), a seguir, completando a segunda fase do algoritmo Backward

$$e_j(n) = d_j(n) - Y_z^j(n). \quad (2.4)$$

onde $d_j(n)$ é o valor do resultado desejado; $Y_z^j(n)$ é o valor do resultado calculado pela rede neural; n é um valor que pertence ao conjunto dos números reais positivos. Então, para finalizar a fase Backward, é calculado o erro e realizada a correção proporcionalmente inversa ao resultado da Expressão do cálculo do erro, ou seja, quanto menor o erro $e_j(n)$, mais treinadas as *sinapses* da rede estarão (VALENÇA, 2009). Desta maneira, pode-se verificar que, para a implementação do algoritmo de retropropagação do erro com aprendizagem supervisionada, faz-se apenas a utilização dos cálculos da Expressão de ativação do neurônio, os cálculos da Expressão de transferência e os cálculos do erro para a correção das *sinapses* e *bias* da rede neural.

2.5.5 Redes Neurais **Kohonen**

À medida que novos estudos foram sendo finalizados, na área de Inteligência Artificial (IA), uma nova técnica de aprendizado foi desenvolvida com uma inovação na realização da distribuição das camadas de neurônios (LUDWIG Jr. e MONTGOMERY, 2007). Este formalismo recebeu o nome de seu criador, proposta pelo professor Teuvo **Kohonen**, em 1982, no trabalho intitulado “*Self-organization formatation of topologically correct feature maps*”. A rede **Kohonen** tem por base um aprendizado não supervisionado, pois foi inspirada nas partes do cérebro humano correspondentes aos comportamentos neurofisiológicos (VALENÇA, 2009). Porém a grande motivação para a criação desta ferramenta, inspirada nas funções simpáticas e parassimpáticas do sistema nervoso autônomo, é a teoria de que o cérebro humano utiliza diferentes regiões específicas do córtex cerebral para entradas sensoriais distintas.

O desenvolvimento desta rede neural por Teuvo **Kohonen** foi considerado extremamente eficiente para a visualização e procura de um conjunto de técnicas e padrões nos dados de entrada com elevadas dimensionalidades. Por isso, nos dias de hoje, ela é largamente utilizada nas áreas que envolvem a Medicina com o tratamento de imagens (ANITHA et al., 2012). As redes **Kohonen**, também conhecidas como mapas autoorganizáveis de **Kohonen** ou SOM's (*self-organization maps*), são muito utilizadas para a mineração de dados (*data mining*) (COSTA e ANDRADE NETO, 2007). Porém, no campo das finanças, ainda não foi totalmente explorada.

Este tipo de rede com aprendizado concorrente tem seu princípio fundamental a competição dos neurônios e clusters, que disputam entre si para ter seus pesos ajustados e responder melhor ao supracitado estímulo (LUDWIG Jr. e MONTGOMERY, 2007). Contudo, um de seus princípios fundamentais consiste em utilizar um processo de cooperação entre os neurônios vencedores que apresentam o ajuste das *sinapses* w_{ij} , e *bias* b_n do neurônio em seus neurônios vizinhos (VALENÇA. 2009).

Para que este ajuste seja empregado corretamente, é necessário aplicar uma função de vizinhança baseada em uma métrica que informe a distância entre o vetor de entrada e o vetor de pesos dos neurônios (LUDWIG Jr. e MONTGOMERY, 2007). Ou seja, é necessário utilizar uma Expressão que remeta à distância como forma de competição entre os neurônios, para ajustar os pesos do neurônio que mais se aproximou do resultado correto e dos neurônios vizinhos a ele.

Para as redes autoorganizáveis, as métricas de aproximação têm um papel muito importante, pois padrões são apresentados em forma de valores de entrada aos mapas de **Kohonen** (COSTA e ANDRADE NETO, 2007). O neurônio vencedor será o mais próximo, de acordo com o critério de similaridade escolhido entre o vetor de entrada e o vetor de *sinapses* w_{ij} da rede.

2.5.5.1 Aprendizado Concorrente Processo Competitivo

Como mencionado anteriormente, o aprendizado não supervisionado da rede **Kohonen** tem por base a competição dos neurônios, que pode ser dividida em duas formas, com três processos cada: o processo competitivo, o processo cooperativo, e o processo adaptativo do vencedor leva-tudo (*winner-take-all*); ou o processo competitivo, o processo cooperativo, e o processo adaptativo do vencedor leva-parte (*winner-take-quota*). No vencedor leva-tudo, apenas os pesos e as *sinapses* do neurônio vencedor se ajustam, já no vencedor leva-parte, os pesos e as *sinapses* do neurônio vencedor e de seus vizinhos se ajustam conforme a métrica de aproximação de similaridade, neste caso, euclidiana, descrita na Expressão (2.5) (VALENÇA, 2009):

$$d_{xw} = \sqrt{\sum_{j=1}^n (x_n - w_{ij})^2}. \quad (2.5)$$

onde d_{xw} é a distância euclidiana do vetor de entrada até o vetor de *sinapses* w_{ij} , e *bias* b_n da rede; x_n são os valores de entrada; n é o número de entradas; e w_{ij} são os pesos sinápticos que ligam a camada de entrada à grade com os neurônios de saída. Imagine-se um vetor de entrada, correspondente por $x = [x_1, x_2, \dots, x_n]^t$, com os valores de x_n selecionados aleatoriamente dentre os demais exemplos de treinamento sem especificar a saída desejada no ciclo t , e um vetor aleatório de pesos $w_{ij} = [w_{i1}, w_{i2}, \dots, w_{in}]^t$, no mesmo ciclo t . O neurônio i , representado pelo valor da *sinapses* w_{ij} , deverá corresponder melhor a esta entrada, ou seja, o vencedor será o que obtiver a distância euclidiana mais próxima a sua entrada, como se pode observar no exemplo a seguir, com as seguintes expressões relacionada a três neurônios:

$$d_{11} = \sqrt{(x_1 - w_{11})^2 + (x_1 - w_{12})^2 + (x_1 - w_{13})^2} \neq x_1;$$

$$d_{12} = \sqrt{(x_2 - w_{21})^2 + (x_2 - w_{22})^2 + (x_2 - w_{23})^2} \cong x_2;$$

$$d_{13} = \sqrt{(x_3 - w_{31})^2 + (x_3 - w_{32})^2 + (x_3 - w_{33})^2} \neq x_3.$$

O neurônio correspondente à distância euclidiana de índice d_{12} , é o vencedor, pois foi o que mais se aproximou do valor de entrada correlativo a ele. A expressão a seguir demonstra este exemplo com valores aleatórios não hipotéticos relacionados ao vetor de entrada $x = [0, 3, 5]^0$, e um vetor de pesos $w_{12} = [1, 2, 3]^0$ será:

$$d_{11} = \sqrt{(0 - 2)^2 + (3 - 2)^2 + (5 - 1)^2} = 21;$$

$$d_{12} = \sqrt{(0 - 1)^2 + (3 - 2)^2 + (5 - 3)^2} = 6;$$

$$d_{13} = \sqrt{(0 - 2)^2 + (3 - 1)^2 + (5 - 1)^2} = 24.$$

Neste exemplo numérico, pode-se melhor observar que o neurônio de índice d_{12} obteve a menor distância euclidiana e o mesmo valor do somatório de todos os pesos *sinápticos* w_{ij} , configurando-se como o neurônio vencedor perfeito. Então, para processos concorrentes (*winner-take-quota*), o neurônio vencedor indica o centro de uma vizinhança topológica de neurônios cooperativos, ou seja, para todos os vizinhos que obtiverem até uma distância lateral permitida do neurônio vencedor, suas *sinapses* também serão corrigidas.

2.5.5.2 Aprendizado Concorrente Processo Cooperativo

À medida que a distância lateral l_{ik} aumenta do neurônio vencedor i em relação aos neurônios visinhos k , decresce suavemente a interação lateral da vizinhança topológica h_{ik} , que indica o grau de cooperação entre o neurônio i e k (LUDWIG Jr. e MONTGOMERY, 2007). Por isso, após o término do processo competitivo e a escolha do neurônio vencedor, utiliza-se o processo cooperativo para definir se os vizinhos do neurônio vencedor também corrigirão seus pesos *sinápticos* (VALENÇA, 2009). Com base na Expressão (2.6), a seguir, é comum o uso da função gaussiana de vizinhança topológica (LUDWIG Jr. e MONTGOMERY, 2007):

$$h_{ik} = e^{\left(-\frac{l_{ik}}{2\sigma^2}\right)} \quad (2.6)$$

onde i refere-se ao neurônio vencedor; k aos neurônios vizinhos; l_{ik} representa a distância euclidiana entre o neurônio i e k ; e é a base natural de logaritmos; e o parâmetro σ é uma qualquer função exponencial que inicia com o mesmo valor do ciclo atual. Para simular o processo de tempo percorrido entre um neurônio e outro, é utilizado o parâmetro σ , ou mais comumente chamando de largura efetiva, que, por sua vez, implica em valores menores ao longo do tempo para o cálculo da distância de interação lateral da vizinhança topológica h_{ik} (VALENÇA, 2009). De maneira geral, para todas as redes neurais, o aprendizado se dá pelo ajuste de seus pesos sinápticos e, nos mapas autoorganizáveis de **Kohonen**, isto não é diferente (LUDWIG Jr. e MONTGOMERY, 2007).

2.5.5.3 Aprendizado Concorrente Processo Adaptativo

Nesta etapa, os pesos do neurônio vencedor e dos demais neurônios que se enquadraram na distância lateral da vizinhança topológica h_{ik} escolhida são ajustados conforme uma métrica de ajuste na direção do vetor de exemplos de entrada realizada pela Expressão (2.7), a seguir (VALENÇA, 2009). Os ajustes serão feitos respeitando a ordem do algoritmo *backpropagation*, ou seja, de traz para a frente, do último neurônio para o primeiro.

$$w_{ij(\text{nov})} = w_{ij(\text{antigo})} + \alpha \cdot h_{ik} \cdot (x_{ij} - w_{ij(\text{antigo})}). \quad (2.7)$$

onde w_{ij} são os pesos de entrada que ligam os exemplos de entrada ao neurônio vencedor; x_{ij} são os exemplos de entrada; h_{ik} é o valor da expressão de interação lateral da vizinhança topológica; e α é a taxa de aprendizagem. No exemplo numérico, a seguir, será utilizada a distância euclidiana d_{11} , calculada no subcapítulo 2.3.5.1, uma taxa de aprendizagem de 0,8 e o resultado da Expressão (2.6) de vizinhança topológica de 0,46, para se poder melhor observar que a correção das *sinapses* do neurônio de índice d_{11} , segunda a Expressão 2.6 será:

$$w_{13(\text{nov})} = 1 + 0,8 * 0,4 * (0 - 1) = 2,47;$$

$$w_{12(\text{nov})} = 2 + 0,8 * 0,4 * (0 - 2) = 2,37;$$

$$w_{11(\text{nov})} = 2 + 0,8 * 0,4 * (0 - 2) = -1,26.$$

Neste exemplo numérico, pode-se observar que os valores corrigidos para os pesos da distância euclidiana d_{11} , são: -1,26 para a *sinapse* w_{11} ; 2,37 para a *sinapse* w_{12} e 2,47 para a *sinapse* w_{13} . Estes valores indicam que o exemplo de entrada está mais próximo do neurônio de saída, como demonstrado, a seguir. Então, corrigindo o neurônio corresponde à distância euclidiana d_{11} , será:

$$d_{11} = \sqrt{(0 - (-1,26))^2 + (3 - 2,37)^2 + (5 - 2,47)^2} = 7,8.$$

Ao término do primeiro ciclo, pode-se verificar que os dois neurônios de saída calculados, d_{11} e d_{12} , estão dispostos a representar os resultados pelos quais foram treinados. Porém, como não foram treinados com um objetivo específico, mas, para melhor observar o funcionamento interno do treinamento de uma rede **Kohonen**, não se pode utilizar estes resultados.

2.5.5.4 Funcionamento Externo da Rede Neural **Kohonen**

Quanto ao funcionamento externo de uma rede neural **Kohonen**, dividem-se os dados de entrada, que são similares entre si, seguindo uma medida de aproximação, no caso a distância euclidiana d_{xw} , agrupando-se os mais próximos, formando classes ou clusters (MILIONI et al., 2007). Cada cluster de dados é definido como uma região no espaço conhecida como grade. A Figura 5, a seguir, demonstra uma grade bidimensional com 2 neurônios de entrada, 4 de neurônio de saída e 2 clusters (a) e (b), onde são definidas as entradas em x_{ij} , as sinapses em w_{ij} , d_{xw} , a distância euclidiana e $w_{ij(novo)}$ são as sinapses corrigidas.

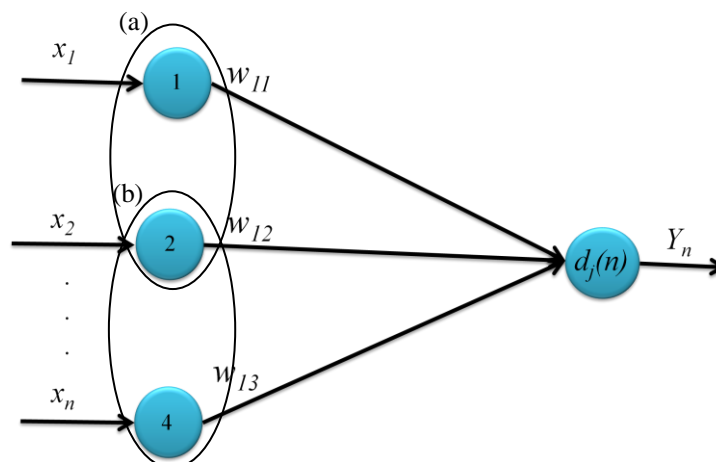


Figura 5 - Topologia de uma Rede **Kohonen**
Fonte: Adaptado de Valença (2009).

Ainda na Figura 5, pode-se observar uma das grandes diferenças entre as redes desenvolvidas, que diz respeito à não disponibilidade de camadas ocultas pelas redes neurais **Kohonen**. Este “problema” inconclusivo em relação a alguma comparação, originalmente transfere todas as desvantagens em relação à não obtenção de camadas intermediárias para as redes neurais com topologias simples, ou seja, somente alguns neurônios de entrada e alguns neurônios de saída.

Segundo Valença (2009), o algoritmo básico SON's de **Kohonen** pode ser resumido nos seguintes passos: no primeiro, no início do processo concorrente do algoritmo, é definida a topologia inicial, a taxa de aprendizagem da rede α , o valor da vizinhança h_{ik} e um valor aceitável de parada a partir do método de tentativa e erro baseado em Sahoo et al., (2012).

No segundo passo, são apresentados os exemplos de entrada para a rede neural, calculando-se a distância euclidiana d_{xw} , no processo concorrente e o valor da vizinhança topológica h_{ik} , no processo cooperativo, correspondentes a todos os neurônios da rede. Ainda no segundo passo, é definido como neurônio vencedor o que obtiver a menor distância euclidiana e/ou o mesmo valor do somatório de todos os pesos *sinápticos* w_{ij} , para fins de determinar que neurônios tenham seus pesos ajustados.

No terceiro passo, no término do processo cooperativo do algoritmo e início do adaptativo, é feito o ajuste dos pesos, ou seja, a definição de novas *sinapses* $w_{ij(novo)}$ para os neurônios vizinhos do neurônio vencedor. Então o neurônio vencedor e todos os neurônios dentro da vizinhança do cluster vencedor terão seus pesos ajustados.

No quarto passo, calcula-se a nova taxa de aprendizado e o valor da vizinhança; no quinto passo, testam-se as condições de parada; no sexto passo, incrementa-se o número de interações e retorna-se ao segundo passo, até que as condições de parada estejam satisfatórias. A Figura 6, a seguir, demonstra o fluxograma com os passos do funcionamento de uma rede **Kohonen**.

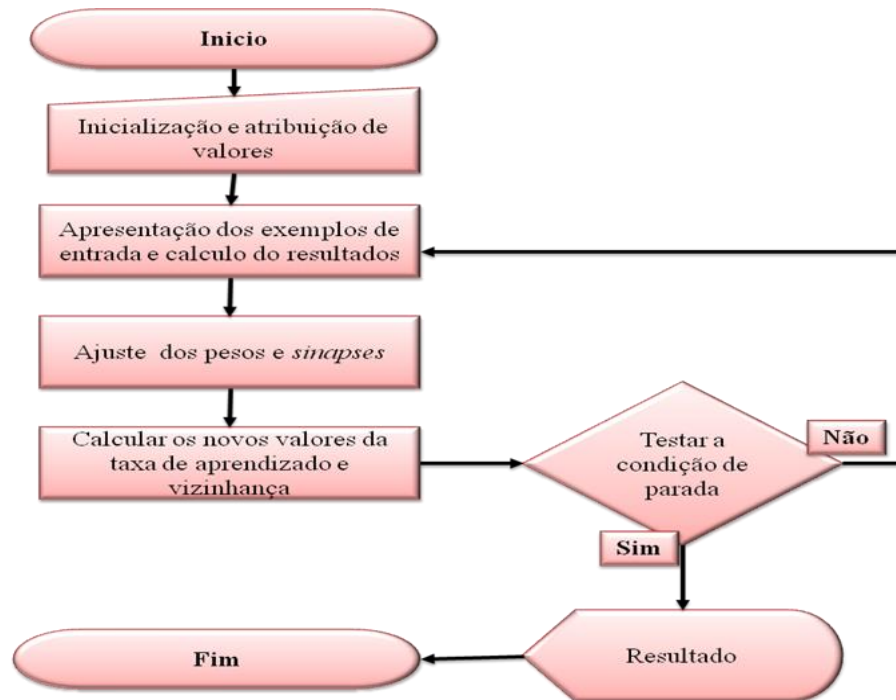


Figura 6 - Fluxograma rede **Kohonen**

Fonte: Organizado pelo autor.

Para a previsão de uma rede neural, é muito importante obter condições de parada ou um erro pequeno, pois, quanto menor, mais próximo da correção das *sinapses* e *bias* ele estará. Então, para que a previsão e a confiabilidade da análise das saídas sejam obtidas, é preciso ter um erro muito pequeno e um R^2 ajustado alto. Mas, para isso, o algoritmo de correção do erro, no caso das redes **Kohonen**, o algoritmo de retropropagação do erro com aprendizagem não supervisionada, deve estar muito bem ajustado.

2.5.5.5 Retropropagação do Erro com Aprendizagem Não Supervisionada

Este método difere da retropropagação do erro com aprendizagem supervisionada pelo simples fato de que não existe um “professor” que avalie a qualidade da saída processada pela rede neural (SOUSA e VALENÇA, 2010). Devido a isso, a rede neural com aprendizado não supervisionada procura padrões estatísticos nas entradas da rede, agrupando-os em classes com representações internas.

Então o professor, neste caso, é responsável pelo fornecimento de padrões não na saída desejada no final da rede, como se viu nas *Perceptrons*, mas no início, junto das entradas de valores. Por isso, este método necessita de mais um neurônio na camada de

entrada para corresponder ao treinamento, como pode ser visto no próximo exemplo numérico.

Para exemplificar o funcionamento das redes **Kohonen**, será utilizado um exemplo de classificação de animais. Neste caso, como mencionado anteriormente, não será fornecida a saída desejada para a rede de tal forma que os dados para o treinamento serão correspondentes à Tabela 4, a seguir.

Tabela 4 – Dados do treinamento da tabela de classificação de animais.

Exemplos	Animal	Glândula	Ovovivíparo	Terrestre	Mamífero (Saída desejada)
1	Galinha	-1	-1	1	0
2	Elefante	1	-1	1	1
3	Peixe	-1	-1	-1	0
4	Ornitorrinco	1	1	-1	1
5	Escorpião	-1	1	1	0
6	Baleia	1	-1	-1	1

Fonte: Valença, 2009.

Neste primeiro exemplo, correspondente ao animal galinha, será treinada uma rede **Kohonen** no último ciclo do treinamento, utilizando uma taxa de aprendizado no valor constante de $0,1\infty = 0,1$. Como demonstrado na Tabela 4, serão analisados dois agrupamentos (Clusters) de animais com o objetivo de classificá-los em mamífero valor 1 e não mamífero valor 0, com valor da vizinhança topológica $h_{ik} = 0$. Para as sinapses, serão utilizados valores de:

$$w_{11} = 0,4765 ; w_{12} = -0,0445 ; w_{13} = 0,1132;$$

$$w_{21} = -0,7638 ; w_{22} = -0,4457 ; w_{23} = 0,1132.$$

Como mencionado no subcapítulo 2.3.5.1, para a primeira abordagem da rede **Kohonen** no neurônio de saída “1”, é necessário calcular a distância euclidiana d_{xw} , levando-se em consideração a Expressão (2.5). Apresentando o primeiro exemplo no ciclo 1, para dois clusters, ou seja, dois neurônios, então, tem-se:

$$d_{11} = \sqrt{(-1 - 0,4765)^2 + (-1 - (-0,0445))^2 + (1 - 0,1132)^2} = 4,0510.$$

Calculando a distância euclidiana para o neurônio de saída “2”, tem-se:

$$d_{21} = \sqrt{(-1 - (-0,7638))^2 + (-1 - (-0,4457))^2 + (1 - 0,1132)^2} = 1,1493.$$

Considerando os dois neurônios, pode-se observar que a menor distância euclidiana foi do neurônio $d_{21} = 1,1493$, que representa o neurônio vencedor. Então, significa que os pesos *sinápticos* w_{ij} do neurônio vencedor estão mais próximos dos valores de entrada x_n . Portanto, os pesos *sinápticos* w_{ij} do vencedor são ajustados conforme a Expressão (2.7) descrita no subcapítulo 2.3.5.3, respeitando a ordem do algoritmo *backpropagation*, ou seja, de traz para frente, do último neurônio para o primeiro.

$$w_{23(novo)} = 0,1132 + 0,1 * (1 - 0,1132) = 0,2019;$$

$$w_{22(novo)} = (-0,4457) + 0,1 * (-1 - (-0,4457)) = -0,5011;$$

$$w_{21(novo)} = (-0,7638) + 0,1 * (-1 - (-0,7638)) = -0,7874.$$

Apresentando-se o segundo exemplo correspondente ao animal elefante, a rede será treinada da mesma forma descrita anteriormente, calculando a distância euclidiana correspondente ao Exemplo 2, elefante, para o neurônio de saída “1”. Assim, seguindo até chegar ao ciclo “3” no último exemplo de treinamento correspondente ao animal baleia, tem-se:

$$d_{11} = \sqrt{(1 - 0,5760)^2 + (-1 - (-0,0260))^2 + (-1 - 0,0072)^2} = 2,1428.$$

Calculando-se a distância euclidiana para o neurônio de saída “2”, tem-se:

$$d_{21} = \sqrt{(1 - (-0,8278))^2 + (-1 - (-0,3959))^2 + (-1 - 0,1735)^2} = 5,0830.$$

Analisando os dois neurônios, pode-se observar que a menor distância euclidiana foi do neurônio $d_{11} = 2,1428$, que se concebe como neurônio vencedor. Então, isso significa que os pesos *sinápticos* w_{ij} do neurônio vencedor estão mais próximos dos valores de entrada x_n . Portanto, os pesos *sinápticos* w_{ij} do vencedor são ajustados conforme a Expressão (2.7) descrita no subcapítulo 2.3.5.3, a seguir, respeitando a ordem do algoritmo *backpropagation*.

$$w_{13(novo)} = 0,0072 + 0,1 * (-1 - 0,0072) = 0,0934;$$

$$w_{12(novo)} = (-0,0260) + 0,1 * (-1 - (-0,0260)) = -0,1234;$$

$$w_{11(novo)} = 0,5760 + 0,1 * (1 - 0,5760) = 0,6184.$$

Ao término do terceiro ciclo do último exemplo, pode-se verificar que o neurônio de saída d_{11} é responsável pelos exemplos correspondentes ao agrupamento dos animais elefante, ornitorrinco e baleia, ou seja, mamíferos. No entanto, o neurônio d_{21} é responsável pelos exemplos correlativos ao agrupamento dos animais galinha, peixe e escorpião, ou seja, não mamíferos. Para testar a capacidade de reconhecimento da rede **Kohonen** treinada anteriormente, serão usados os últimos pesos *sinápticos* $w_{ij(novo)}$ ajustados e outros exemplos na Tabela 5, a seguir, que não foram empregados no treinamento.

Tabela 5 – Dados da tabela de classificação de animais.

Exemplos	Animal	Glândula	Ovovivíparo	Terrestre	Mamífero (Saída desejada)
7	Equidna (Ouriço)	1	1	1	-
8	Anaconda	-1	1	-1	-

Fonte: Valença, 2009.

No primeiro exemplo, após o treinamento correspondente ao animal equidna, a rede será utilizada da mesma forma descrita anteriormente, calculando a distância euclidiana correspondente ao Exemplo 2, equidna, para os dois neurônios d_{11} e d_{12} , mas pulando a fase do processo adaptativo. Apresentando o primeiro exemplo no ciclo 1, para dois clusters, ou seja, dois neurônios, tem-se:

$$d_{11} = \sqrt{(1 - 0,6184)^2 + (1 - (-0,1234))^2 + (1 - 0,0934)^2} = 2,6035.$$

Calculando a distância euclidiana para o neurônio de saída “2”, tem-se:

$$d_{21} = \sqrt{(1 - (-0,8278))^2 + (1 - (-0,3959))^2 + (1 - 0,1735)^2} = 5,9725.$$

Analisando os dois neurônios, pode-se observar que a menor distância euclidiana foi do neurônio $d_{11} = 2,6035$, que representa o neurônio vencedor. Então, significa que os pesos *sinápticos* w_{ij} do neurônio vencedor estão mais próximos dos valores de entrada x_n . Portanto, este animal será agrupado corretamente no grupo 1, chamado de grupo dos mamíferos. Analisando-se a anaconda, o segundo animal após o treinamento e usados os últimos pesos *sinápticos* $w_{ij(novo)}$ ajustados, tem-se:

$$d_{11} = \sqrt{(-1 - 0,6184)^2 + (1 - (-0,1234))^2 + (-1 - 0,0934)^2} = 4,7032.$$

Calculando a distância euclidiana para o neurônio de saída “2”, tem-se:

$$d_{21} = \sqrt{(-1 - (-0,8278))^2 + (1 - (-0,3959))^2 + (-1 - 0,1735)^2} = 3,3554.$$

A distância euclidiana representada pelo neurônio $d_{12} = 3,3554$ foi a menor dentre as duas testadas, que o concebe como neurônio vencedor. Isso significa que os pesos *sinápticos* w_{ij} do neurônio vencedor estão mais próximos dos valores de entrada x_n . Logo, o animal anaconda será agrupado corretamente no grupo 2, chamado de grupo dos não mamíferos.

Como se pode observar no exemplo anterior, esta abordagem não faz uso de nenhuma intervenção humana na escolha dos padrões escolhidos. Portanto não se utiliza da saída desejada para o ajuste do erro dos pesos *sinápticos* w_{ij} como nas redes **Perceptron**, porém utiliza-se da distância euclidiana, descrita na Expressão (2.5), para alcançar o resultado esperado, caracterizando-o como um método não supervisionado (VALENÇA, 2009).

Tanto as redes neurais **Perceptron** como as **Kohonen** utilizam-se do algoritmo *backpropagation* de correção das *sinapses*. Como mencionado anteriormente, este algoritmo utiliza-se da retrocorreção dos pesos *sinápticos*, da camada de saída para as camadas intermediárias (BARBOSA et al., 2005), descrita no subcapítulo 2.3.4.2.2, para as **Perceptron** e no subcapítulo 2.3.5.3, para as redes **Kohonen**.

Porém o algoritmo *backpropagation* é somente utilizado no treinamento / aprendizagem de uma rede neural, pois, nos próximos passos do desenvolvimento de uma rede, não serão necessárias correções *sinápticas*, visto que, se a rede não obtiver um resultado satisfatório, será indispensável refazer todos os passos anteriores.

2.5.6 Treinamento / Aprendizagem, Validação Cruzada e Verificação das Redes Neurais

Segundo Ludwig Jr. e Montgomery (2007), para que uma rede neural tenha credibilidade em seus resultados, ela necessita de três passos sequenciais da implementação da rede neural: o primeiro passo é o desenvolvimento correto das redes neurais utilizadas descritos nos subcapítulos 2.3.4.2 e 2.3.5; o segundo passo é a implementação da ordem dos cálculos dentro do neurônio descrito no subcapítulo 2.3.4.2.1, e subcapítulos 2.3.5.1, 2.3.5.2; e o terceiro e último passo do desenvolvimento diz respeito ao ajuste dos *bias* e *sinapses* do neurônio no treinamento / aprendizado descrito nos subcapítulos 2.3.4.2 e 2.3.5.3.

Após o desenvolvimento de uma rede neural, há três etapas chamadas de treinamento que fornecem a calibração dos parâmetros e a definição do ponto ótimo de parada da rede: o treinamento / aprendizado, a validação cruzada e a verificação (VALENÇA e LUDERMIR, 2007; SAHOO et al., 2012). Para melhor entendimento, somente o treinamento / aprendizado será descrito novamente, de uma forma mais simplista.

2.5.6.1 O Treinamento / Aprendizado

O treinamento / aprendizado da rede neural *Perceptron* de múltiplas camadas (MLP) leva em consideração uma regra formal, a utilização de informações do supervisor da rede. Este supervisor atualiza os parâmetros modificáveis da rede neural, como os *bias*, a taxa de aprendizado, o número de ciclos, o valor do erro e as *sinapses* para chegar a um resultado aceitável de erro. A participação do supervisor pode ser classificada em métodos de treinamento / aprendizado forte e fraco.

Como se expôs anteriormente, quando a participação for forte, o supervisor fornece um conjunto de valores de entradas correlacionados com um conjunto de valores de saída desejadas na camada de saída. Este método corresponde ao algoritmo *backpropagation*, onde o erro é retropropagado da saída para a entrada, descrito anteriormente nos subcapítulos 2.3.4.2, fazendo com que a rede “imite” o supervisor (NASCIMENTO Jr. e YONEYANA, 2004; SCHALKOFF, 2001).

No caso das redes neurais não supervisionadas como a **Kohonen**, o treinamento / aprendizado não levará em consideração o erro da comparação entre os parâmetros saída desejada e saída calculada para ajuste dos pesos. Neste tipo de rede, diferente das redes *Perceptron*, a informação de treinamento / aprendizado corresponde à distância entre as

entradas x_n e as sinapses w_{ij} da rede neural e a outro neurônio na camada de entrada, ou seja, as redes neurais **Kohonen** ajustam-se aos pesos *sinápticos* do neurônio conforme uma métrica de aproximação de similaridade ou padrões.

Em uma rede **Kohonen**, essa correção entre os parâmetros desejados e a real saída da rede neural é feita com base em um algoritmo de seleção de neurônios, que seleciona os neurônios vencedores a partir da menor distância euclidiana entre as entradas x_n e as sinapses w_{ij} como demonstrado no subcapítulo 2.3.5.5. Como mencionado anteriormente, este método de correção corresponde a duas formas distintas de ajuste de *sinapses* dos *clusters* vencedores: o vencedor leva-tudo (*winner-take-all*) e o vencedor leva-parte (*winner-take-quota*) (VALENÇA, 2009) descritas no subcapítulo 2.3.5.1.

2.5.6.2 A Validação Cruzada

Já foi dito que, após a conclusão do treinamento / aprendizado da rede neural, é necessário o desenvolvimento da segunda etapa do treinamento, a etapa da validação cruzada. Esta etapa, considerada como a segunda mais importante para o treinamento, expõe precisamente quando a rede neural estará treinada para uma nova série de dados sem o *overfitting* (treinamento em excesso) (VALENÇA, 2009).

Então, a validação cruzada pode ser caracterizada como um critério de parada para o treinamento / aprendizado, pois, nesta etapa, descobre-se se a rede neural não está super treinada para a entrada correspondente (VALENÇA e LUDERMIR, 2007), visto que a rede neural não deverá ser exclusivamente treinada, na etapa do treinamento / aprendizado, para somente uma série de dados e sim ser tão generalizável quanto possível.

Para a identificação do momento certo de parada do treinamento, na validação cruzada, primeiramente busca-se a minimização do erro de três expressões mais utilizadas e, posteriormente, a comparação dos resultados com o erro do treinamento / aprendizado das mesma três expressões, que são o Erro Médio Quadrado, apresentado na Expressão (2.8); o Erro Médio Global, apresentado na Expressão (2.9) (LUDWIG Jr. e MONTGOMERY, 2007); e o Erro Padrão de Predição, apresentado na Expressão (2.10), a seguir. Na minimização do erro com estas equações, procura-se um resultado aceitável de 0,08 a 0,05 por ciclo no Erro Médio Global, um resultado de 0,008 a 0,005 por ciclo no Erro Médio Quadrático e um valor de 0,008 a 0,005 por ciclo no Erro Padrão de Predição (VALENÇA, 2009; GUIMARÃES, 2008).

$$EMQ = \frac{1}{N} \sum_{n=1}^N (Y_j(n) - d_j(n))^2; \quad (2.8)$$

$$EMG = \frac{1}{N} \sum_{n=1}^N Y_j(n) - d_j(n); \quad (2.9)$$

$$EP = \left[\frac{1}{N} \sum_{n=1}^N (Y_j(n) - d_j(n))^2 \right]^{0,5}. \quad (2.10)$$

onde $Y_j(n)$ é o valor calculado pela rede neural; $d_j(n)$ é o valor desejado pelo supervisor; N é o número total de valores do conjunto de verificação, tal que N é um valor IR^+ ; e n é o número atual do valor da entrada. Na segunda parte da validação cruzada, a comparação do erro das três expressões nas duas etapas do treinamento, busca-se localizar o ponto de parada ideal, quando o valor mínimo do erro na validação cruzada começa a aumentar, mesmo que o erro do conjunto de treinamento / aprendizado continue a diminuir (VALENÇA, 2009). A Figura 7 demonstra um gráfico com o ponto de parada ideal para o treinamento de uma rede neural fictícia. Então, a partir da identificação do ponto de parada ideal do treinamento, a rede neural é considerada treinada e validada até o ciclo que corresponde ao ponto de parada ideal demonstrado na Figura 7, a seguir.

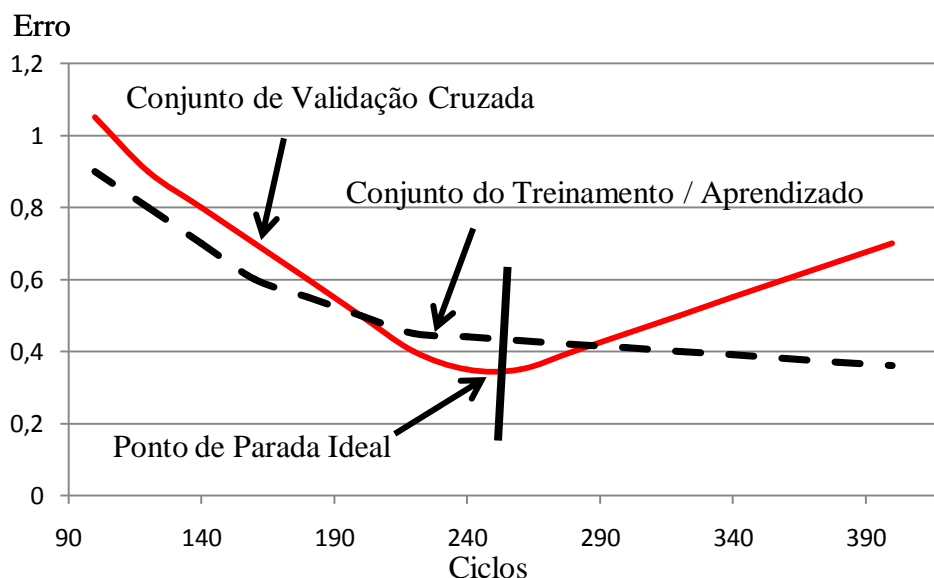


Figura 7 – Gráfico do Ponto de Parada Ideal.

Fonte: Valença, 2009.

2.5.6.3 A Verificação

O último quesito a ser avaliado no treinamento é a Verificação, onde há uma comparação entre o resultado do erro médio quadrado da verificação com o erro médio quadrado da validação cruzada. Neste estágio do treinamento de uma rede neural, é possível avaliar o desempenho do modelo utilizado e obter o ponto exato em que a rede neural deixa de estar treinada. A partir do momento em que as entradas x_n , da rede neural na Validação Cruzada resultam do Erro Médio Quadrático afastado do Erro Médio Quadrático da Verificação, são descartados os valores seguintes de entradas x_n da rede neural, como mostra a Figura 8, a seguir.

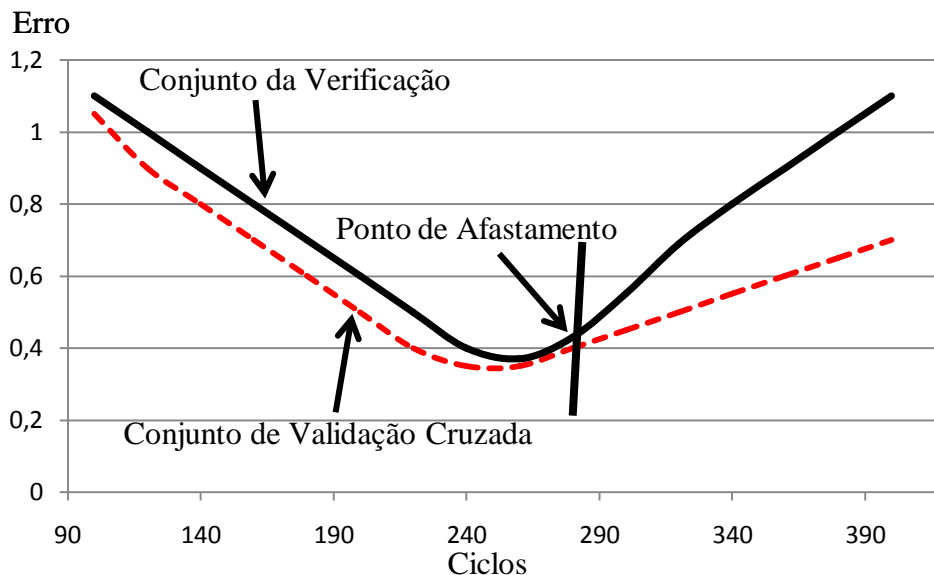


Figura 8 – Gráfico do Ponto de Afastamento.

Fonte: Organizada pelos autores.

Posteriormente a esta avaliação entre o erro da verificação e o erro da validação cruzada, a rede neural é considerada treinada, validada e verificada até o exato momento em que os erros se afastam. Com esta técnica de avaliação do treinamento, a rede neural demonstra que ainda consegue reconhecer as características ou padrões pelos quais foi treinada até o ponto do afastamento (VALENÇA, 2009). Porém, para a comparação entre os métodos, geralmente são utilizadas outras expressões descritas, a seguir, na verificação entre os métodos.

2.5.6.4 A Verificação Entre os Métodos

Após a comparação entre as etapas do treinamento, pode-se fazer a primeira comparação da previsão dos dois métodos utilizados. O critério da comparação será o coeficiente de determinação ajustado ou corrigido pelos graus de liberdade, mais comumente chamado de R^2 ajustado.

Para o funcionamento do R^2 ajustado, é utilizado o percentual de aproximação de um determinado valor estabelecido. Portanto, quanto mais próximo do valor 1 o R^2 ajustado se encontrar, de acordo com a Expressão 2.11, a seguir, mais garantia de previsibilidade correta o método correspondente terá. Segundo Alves (2007) e Lopes et al., (2008), quanto maior o R^2 ajustado, maior será o poder de explicação de X em Y e mais bem treinado estará o método.

$$\overline{R^2} = 1 - \frac{t_o - 1}{t_o - T_p} (1 - R^2). \quad (2.11)$$

onde T_o é o número total de observações; T_p é o número de variáveis ou o número de neurônios na camada de entrada; R^2 é o coeficiente de determinação; e $\overline{R^2}$ é o coeficiente de determinação ajustado. Para o segundo critério de comparação entre os métodos, será utilizado o erro médio quadrático e o erro médio global da etapa da Verificação, descrito na Expressão 2.8.

2.5.7 Topologias das Redes Neurais

Foi referido anteriormente que uma rede neural é um conjunto de neurônios predispostos em uma sequência que recebe o nome de topologia de uma rede neural. Essa série é composta pelo somatório do número de neurônio nas camadas de entrada N_{inp} , camadas escondidas (no caso das redes *Perceptrons*) N_{hid} e camada de saída N_{out} Expressão (2.12), a seguir.

$$TP_p = \sum_{p=1} N_{inp} ; TP_e = \sum_{e=1} N_{hid} ; TP_s = \sum_{s=1} N_{out} . \quad (2.12)$$

Para a escolha da topologia, é feita uma comparação com base no erro médio quadrático, delinear na Expressão (2.8), com outras topologias testadas no treinamento da rede neural e então utiliza-se a adaptação de neurônio que melhor se enquadrou ao problema e a que obtiver menor erro, ou seja, é utilizada a sequência de neurônios que melhor compreendeu ao problema demonstrado no treinamento. Mas, para fins de teste, este enquadramento somente é verificado no treinamento da rede neural, com o método de estimativa pelo conjunto de pesos a serem treinados descrito em Valença, (2009), ou o método de tentativa e erro descrito em Sahoo et al., (2012).

Como o objetivo da comparação de topologias consiste em expor qual se adere melhor ao plano de resultados almejados, levando-se em consideração as entradas, as saídas e o tipo de rede neural utilizado, este é um ponto muito importante na análise dos resultados (VALENÇA, 2009; LUDWIG Jr. e MONTGOMERY, 2007; GONÇALVES, 2005). Por exemplo, redes com a topologia 2-3-3, como pode ser visto na Figura 9, a seguir, com duas entradas, três neurônios na camada oculta e três neurônios na camada de saída.

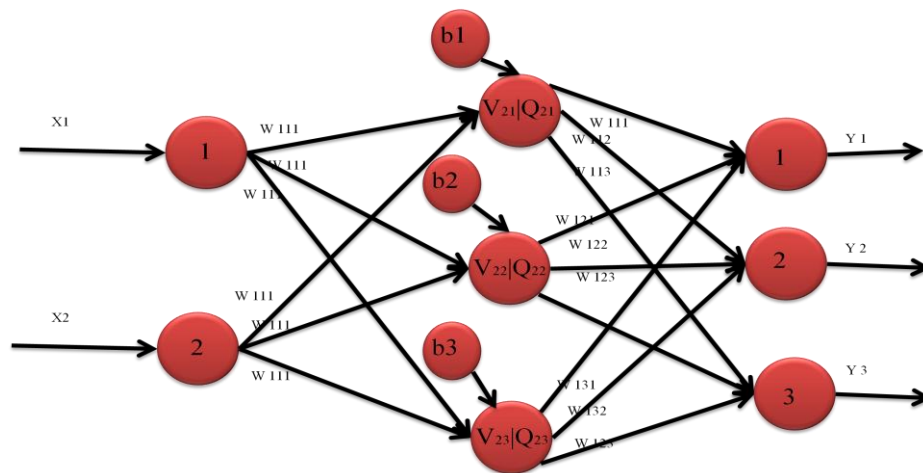


Figura 9 - Topologia Rede Neural Artificial MLP

Fonte: Organizada pelo autor

2.5.8 Iluminando a “Caixa Preta”

Embora as redes neurais contenham um poder de explicação melhor, na maioria dos casos, do que muitos modelos tradicionais citados na literatura como em Ceretta, et al., (2010), Cunha et al., (2010), Oliveira et al., (2010) e Li e Fu, (2012), ainda vários pesquisadores das áreas financeiras relutam em utilizá-las, pois estes modelos são

considerados “caixas pretas” em relação a pouca explicação das variáveis independentes com o processo de predição.

Entretanto, esta crítica ocorre pelo fato de que a explicação entre as variáveis de entrada (independente) e as de saída (dependente) são difíceis de serem compreendidas somente por meio das *sinapses* da rede neural. Em outras palavras, a compreensão do mecanismo de introspecção subjacente das inter-relações entre as variáveis de saída e de entrada de uma rede neural não são fornecidas simplesmente pelas *sinapses* (VALENÇA e LUDERMIR, 2007; VALENÇA, 2009).

Portanto, o relacionamento entre redes neurais e o conceito de “caixa preta” está intrinsecamente relacionado à importância das variáveis e neurônios de entrada para a previsão. Ou seja, se uma variável de entrada ou neurônio contiver pouca influência na saída do resultado de uma rede, não teria necessidade de mantê-la no modelo (SOUSA e VALENÇA, 2010).

A utilização do modelo sugerido por Garson (1991) modificado por Goh (1995), por conseguinte, determinará a influência total e a eliminação dos neurônios das variáveis dependentes, com pouca ou nenhuma influência nas variáveis independentes. O modelo de Goh (1995) permite a utilização em algoritmos de correção do erro como os *backpropagation* com base em derivadas parciais da saída com relação aos pesos nas conexões durante o treinamento da rede neural descrita no método do trabalho (VALENÇA e LUDERMIR, 2007).

2.6 A Teoria das ondas de Elliott

Ralph Nelson Elliott foi um contador americano, considerado por muitos um gênio, que, por volta dos anos de 1938, na sua monografia denominada de “O princípio das ondas de Elliott”, desenvolveu uma ferramenta de análise técnica para auxiliar na identificação das tendências do mercado acionário. Nesse trabalho, afirmava que o universo era regido por leis sem as quais haveria o caos, e onde há o caos, não há nada. Porém foi em 1946, no trabalho intitulado “*Natures Law: The Secret of the Universe*” ou “Lei da Natureza: O Segredo do Universo”, que escreveu seu trabalho definitivo sobre os padrões das ondas do mercado acionário (NORONHA, 2009).

Nestas condições, Elliott classificava o mercado de ações como participante de uma lei natural que governa todas as atividades humanas, ou seja, a lei do comportamento psicológico das massas que apresenta o perfil da população, especificando características culturais e

raciais nos movimentos das atividades humanas (CHAVES, 2003; ELLIOTT, 2011). Para ele, o conhecimento desta lei, assim como os padrões relacionados a ela, quando aplicados ao mercado acionário, poderiam fazer com que a previsão de tais comportamentos fosse possível a longo prazo, utilizando o preço e/ou o volume das ações.

Então, em sua teoria das ondas, Ralph Nelson Elliott afirma que o mercado acionário é composto por padrões de movimentos repetitivos que se modificam à medida que a lei natural que governa todas as atividades humanas se adapta para outras tendências psicológicas (HAYASHI, 2002). No mercado de ações, esses comportamentos psicológicos das massas expressam-se por meio de repetitivos cinco padrões de ondas a favor da tendência, seguidas de três ondas de correção para cada tendência (NORONHA, 2009). Um diagrama perfeito do ciclo do mercado de ações segundo Elliott está exposto, a seguir, na Figura 10.

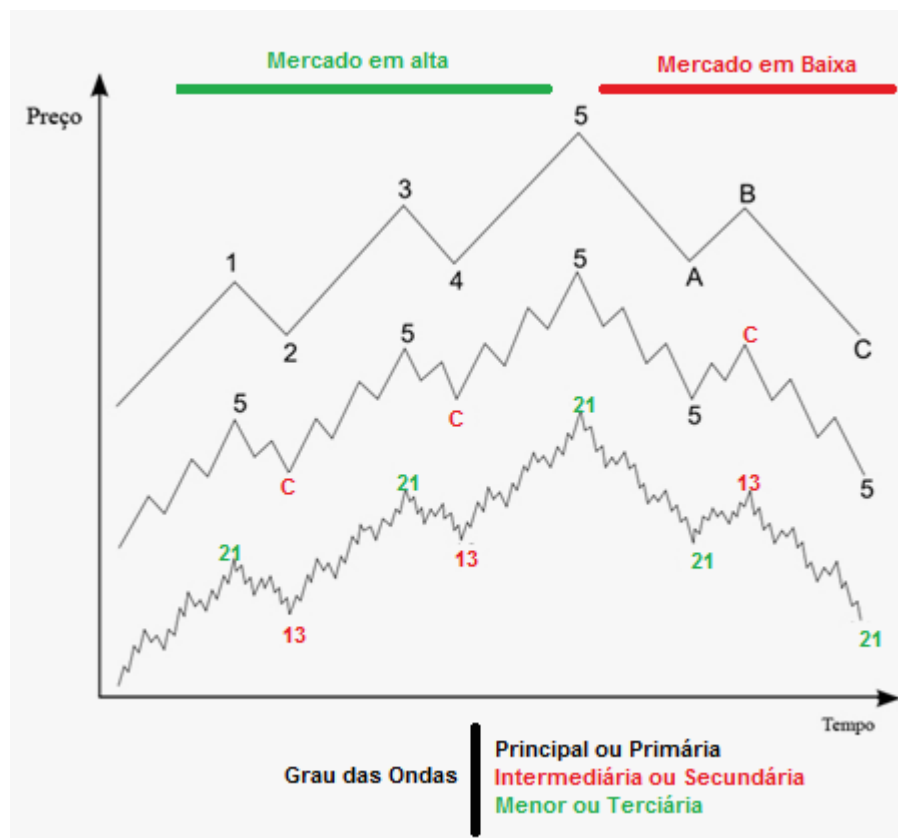


Figura 10 - Ciclo do Mercado de Ações acompanhado das Ondas de Elliott.

Fonte: Noronha (2009).

Na Figura 10 acima, pode-se observar que a conclusão inicial dos padrões do mercado de ações é dividida em “Mercado de Alta”, com cinco ondas principais designadas pela numeração 1, 2, 3, 4, e 5, e “Mercado de Baixa”, com três ondas principais designadas pelas letras A, B, e C. Ondas 1, 3 e 5 são caracterizadas como impulsoras a favor da tendência e as ondas 2 e 4 são definidas como de correção de suas antecessoras, ou seja, a onda 2 corrige o

excesso da onda 1 e a onda 4 corrige o excesso da onda 3 (NORONHA, 2009). Já nas ondas do mercado de baixa, pode-se classificá-las em impulsoras a favor da tendência de baixa A e C e a onda B, de correção da onda A.

Então, voltando-se à conclusão inicial de que o ciclo inteiro do mercado é dividido em mercado de alta e mercado de baixa, pode-se afirmar que as cinco ondas primárias são divididas em 21 ondas secundárias e assim, sucessivamente. Já as três ondas do mercado de baixa primárias são divididas em 13 ondas secundárias e assim, sucessivamente.

Na visão de Elliott, esta relação da utilização de somente números que fazem parte da sequência de Fibonacci [3, 5, 13, 21, ...], não é mera coincidência. Para ele, os movimentos do mercado seguem a lei natural que governa todas as atividades humanas que, por consequência, seguem uma sequência de Fibonacci. Ou seja, o mercado de ações expressa-se em uma razão aritmética que compreende a sequência de Fibonacci denominada de o número de ouro. Este número relaciona-se com a razão da sequência de Fibonacci, que pode ser de 0,618 para o número posterior e 1,618 para o número antecessor, a um valor da sequência de Fibonacci.

Como pode ser observado, a sequência de Fibonacci compreende todos os números sugeridos desde o início do parágrafo anterior. A título de curiosidade, relacionando o corpo humano com os números 3 e 5 encontrados na sequência de Fibonacci (NORONHA, 2009), pode-se concluir que o nosso dorso tem cinco extensões: uma cabeça; dois braços; e duas pernas. Cada perna e cada braço são subdivididos em três seções. Os braços e as pernas terminam em cinco dedos. Os dedos da mão e do pé, exceto o dedão, são divididos em três seções. A sequência de Fibonacci também pode ser encontrada em várias outras relações, sendo até considerada como a digital do universo.

Estas relações entre a sequência de Fibonacci podem ser encontradas nas teorias já estabelecidas de configurações biológicas. Como exemplo, a disposição dos galhos das árvores ou das folhas em uma haste, que crescem de acordo com a sequência de Fibonacci (DOUADY e COUDER, 1995). Em relação à mesma proporção da espiral da sequência de Fibonacci, pode ser citado o arranjo do cone da alcachofra, do abacaxi (JONES e WILSON, 2006) ou o desenrolar de uma samambaia.

A lista de fenômenos relacionados com a sequência de Fibonacci é ainda maior e poder-se-ia continuar citando uma infinidade deles. Porém, o objetivo aqui é de apenas dar uma pequena ideia da importância da sequência de Fibonacci para os padrões do universo e da teoria das ondas.

Entretanto, apesar da baixa aceitação da relação da sequência de Fibonacci com o mercado acionário entre os pesquisadores das áreas financeiras, baseada em somente trabalhos empíricos, a teoria das ondas de Elliott é considerada por vários estudiosos das áreas de análise técnica como a que mais se aproxima das oscilações de tendência do mercado, pois foi totalmente elaborada com base em observações de índices de preços (NORONHA, 2009). Esta técnica foi desenvolvida a partir da identificação de padrões de movimentos, sem modificar os preços (HAYASHI, 2002), de modo semelhante às ideias de Huyghebaert e Wang (2010) e Khalid e Rajaguru (2007), os quais afirmam que o mercado de renda variável seria também composto por padrões de movimentos que se modificam a partir de um acontecimento marcante.

2.6.1 A Base Matemática da Teoria

Segundo Elliott, todo o comportamento humano pode ser previsto no mercado de ações, se for mapeado com base em dois aspectos distintos, padrões e razões, e todos os dois se expressam por meio de ondas correspondentes a uma sequência de Fibonacci. Na teoria das ondas de Elliott, todos os padrões de ondas são comuns, mas não constantes e podem ser visualizados a partir do padrão que o precede.

Dessa forma, para a classificação de uma onda qualquer, é necessário desvendar o início do ciclo, ou seja, a primeira onda. Porém, pelo fato de o mercado acionário não ser composto por ciclos de padrões perfeitos, será necessário utilizar basicamente nove regras listadas por Elliott, que tornam a identificação das ondas possível (NORONHA, 2009).

A primeira regra diz respeito à terceira lei de Newton, pois, no mercado acionário, toda ação provoca uma reação na mesma direção e em sentido contrário. A segunda regra de Elliott relata que o ciclo inteiro do mercado é composto por oito ondas, cinco a favor da tendência, seguidas de três ondas de correção. A terceira regra diz respeito à identificação da primeira onda, que geralmente inicia-se em um fundo³ importante para o mercado. A Figura 11, a seguir, demonstra vários fundos que podem ser utilizados ao longo da história do Ibovespa.

³ Fundo é o ponto mais baixo de um ciclo de cotações em um gráfico (DEBASTIANI, 2008).

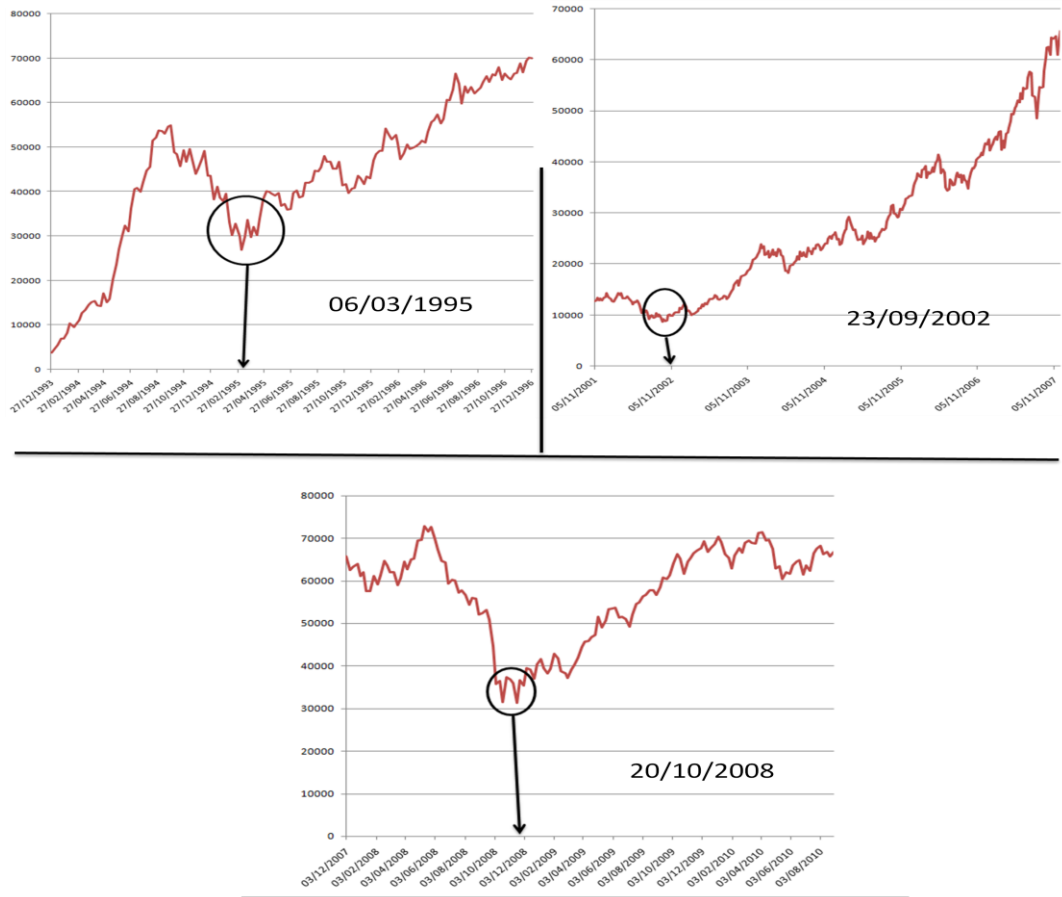


Figura 11 – Exemplos da terceira regra da Teoria das Ondas de Elliott

Fonte: Organizada pelos autores.

Na Figura 11, podem-se observar três gráficos, contendo cada um, um fundo demarcado com um círculo. Estes fundos considerados importantes são denominados de início de ciclo, ou seja, início da onda um. A quarta regra propõe que a periodicidade das ondas ímpares não muda o padrão, ou seja, as ondas um, três e cinco podem ser esticadas ou comprimidas, mas o padrão intrínseco de oito ondas é constante. A quinta regra relaciona-se com a ocorrência de prováveis extensões em uma das ondas ímpares, ou seja, movimentos exagerados que podem aparecer em uma das três ondas, como mostra a Figura 9, a seguir.

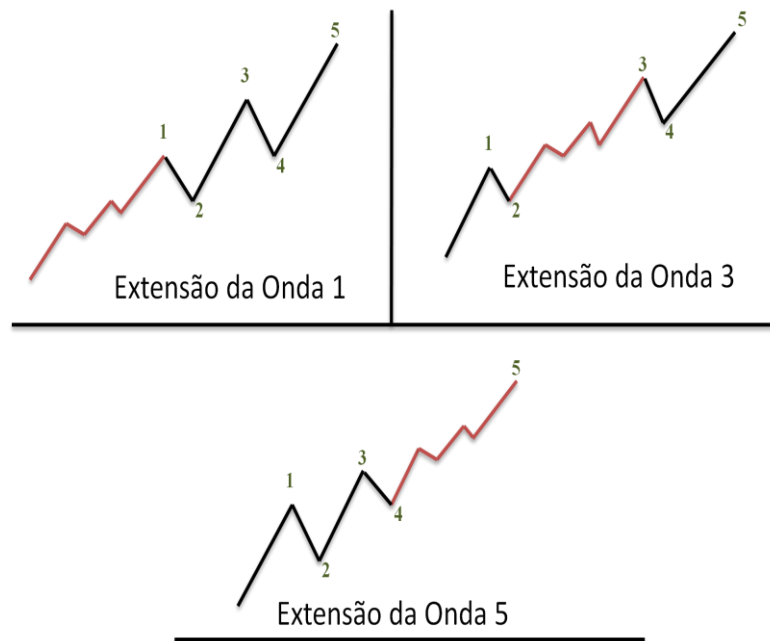


Figura 12 – Extensão das Ondas de Impulso

Fonte: Noronha, 2009.

Na Figura 12, podem-se ressaltar três gráficos contendo cada uma das três extensões de impulso possíveis. Na extensão da onda um, pode-se observar que a onda um foi esticada. Na extensão da onda três, pode-se observar que a onda três foi esticada e, na extensão da onda cinco, que a onda cinco foi esticada.

A sexta regra da teoria das ondas de Elliott refere-se à identificação da segunda onda que, em 73% das vezes, corrige de 50 a 62% do tamanho da onda um, ou seja, a razão da onda dois, mais comum, é ter de 50 a 62% do tamanho da onda um em uma tendência contrária à da principal. A sétima regra faz referência à identificação da terceira onda e relata que a onda três tem razão de 1,618 (numero de ouro) até 1,75 vezes o tamanho da onda um. A oitava regra da teoria das ondas de Elliott relata que, em um movimento de alta, o fundo da onda quatro não poderá ultrapassar o topo⁴ da onda um. A Figura 13, a seguir, demonstra esta relação.

⁴ Topo é o ponto mais alto de um ciclo de cotações em um gráfico (DEBASTIANI, 2008).

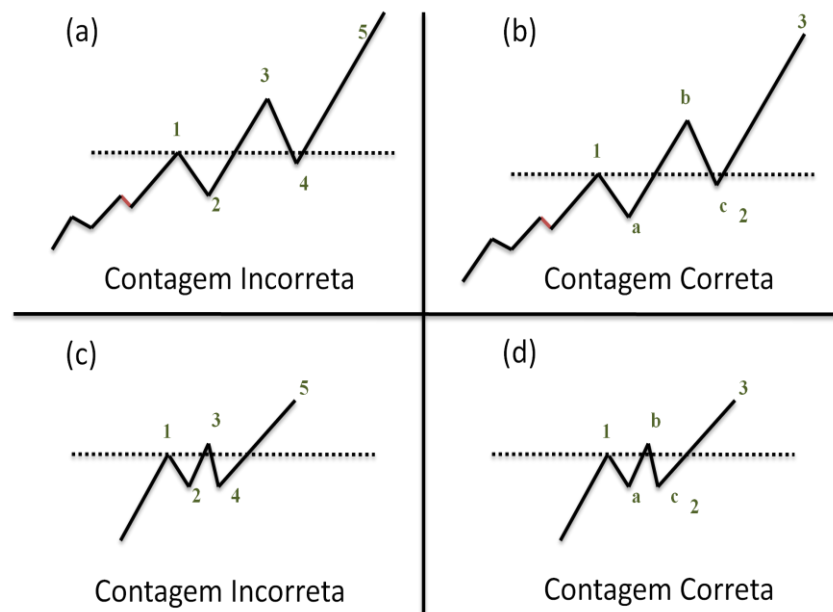


Figura 13 – Contagem da onda quatro, oitava regra da teoria das ondas.

Fonte: Organizada pelos autores.

Na Figura 13 (a) e 13 (b), mais para fins explicativos, pode-se observar um gráfico com duas contagens de ondas diferentes, uma correta e uma incorreta, assim como na Figura 13 (c) e 13 (d). A nona regra da teoria das ondas de Elliott faz referência à identificação do final da quinta onda, onde é usualmente utilizado o desenho de uma cunha ou triângulo como mostra a Figura 14, a seguir.

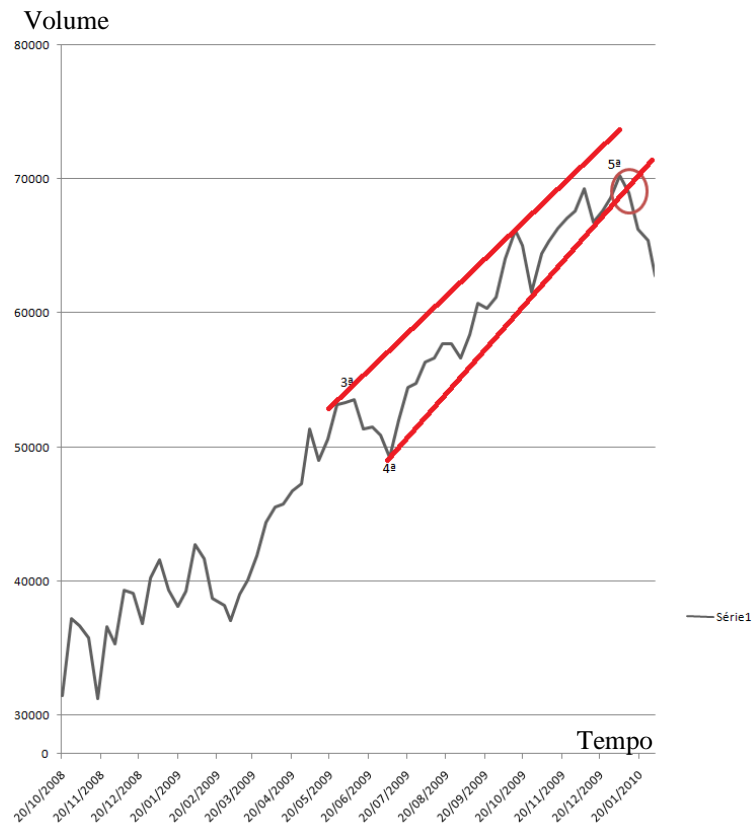


Figura 14 – Triângulo da nona regra da Teoria das Ondas de Elliott

Fonte: Organizada pelos autores.

Na Figura 14, pode-se observar a aplicação da nona regra, que diz respeito ao término da quinta onda em um gráfico com as cotações do Ibovespa de 20/10/2008 até 20/01/2010. Na medida em que o triângulo é rompido, demonstra-se o fechamento do ciclo da quinta onda. Para melhor identificação do *timing* exato, está circulado o ponto preciso no gráfico do término da quinta onda e o começo das três ondas de correção. Ao término da aplicação da nona regra, pode-se dar por finalizado o ciclo e o final da contagem das ondas de impulso (Mercado de Alta). Então se deve acrescentar, enfim, que, com base no seguimento rigoroso destas nove regras, pode-se obter uma contagem correta das ondas.

Entretanto, como pode ser observado nas nove regras descritas anteriormente, a maioria delas apresenta muito espaços para grandes interpretações dos gráficos em relação ao seu uso. É possível ainda acrescentar que, apesar da usabilidade criteriosa destas regras, elas apresentam uma variação que pode ser encontrada em Noronha (2009), Chaves (2003), Elliott (2011) ou Debastiani (2008), garantindo ainda mais a contagem ponderada de cada onda.

3 MATERIAL E MÉTODOS

Neste capítulo e nos subcapítulos seguintes, serão apresentadas as técnicas e os testes estabelecidos para a modelagem e aquisição das entradas no treinamento / aprendizado (calibração dos parâmetros modificáveis), validação cruzada (definição do ponto ótimo de parada), verificação (validação dos modelos). Também será apresentado, no subcapítulo 3.2, como se dará a obtenção dos resultados dos métodos heurísticos utilizados, assim como a tentativa de previsão de uma crise econômica com o modelo que mais se destacou no conjunto das etapas do treinamento.

Para um melhor entendimento dos métodos utilizados, este capítulo será dividido em três etapas: manipulação dos dados, utilizando fórmulas e testes; métodos de treinamento, demonstrando como se dará a aquisição do conhecimento nas duas redes neurais; e os testes utilizados, que verificarão se a série utilizada está de acordo com os estudos científicos apresentados no referencial teórico.

3.1 Manipulação dos Dados

Após o desenvolvimento das ferramentas heurísticas, para que a modelagem da série do retorno do volume de fechamento semanal do Ibovespa esteja de acordo com estudos empíricos e científicos realizados na área, é preciso seguir três grandes passos: o primeiro consiste em dividir a série nas etapas do treinamento para as duas redes neurais; o segundo passo consiste em normalizar a série para que as redes deem a mesma importância a todas as entradas; e o terceiro requer transformar a série utilizada para uma série de Raiz Unitária, empregando a expressão de defasagem de x_n , ou seja, retorno.

Os retornos do mercado de capitais são indicadores que refletem os graus de pressão especulativa, independente da própria decisão política monetária de um país (BARBA, 2010). Autores como Rigobon (2002); Didier (2010) defendem que quase todas as pressões especulativas e crises financeiras não afetam indicadores como taxas de juros ou gastos em investimentos, tanto quanto os retornos dos volumes dos mercados de capitais.

Dessa forma, para satisfazer as afirmações de Barba (2010), Rigobon (2002) e Didier (2010), será utilizado o retorno do volume do Ibovespa como entrada de valores nos métodos. Após a divisão dos valores de entrada dos métodos em duas partes, antes e depois da crise de 2008, estes valores ainda serão divididos em quatro categorias: saída desejada, Treinamento / Aprendizado, Validação Cruzada e Verificação.

A primeira divisão dos valores de entrada, segundo Huyghebaert e Wang (2010), deverá ser em duas partes para levar em consideração que os padrões do comportamento dos preços e volumes do mercado são diferentes, antes, durante e depois de uma crise mundial. Para o lag⁵ das redes neurais será utilizado o primeiro trimestre de 2007, pois foi o começo do crescimento de forma abrangente da inadimplência do *Subprime* e o não pagamento dos empréstimos (SOARES et al., 2009). Portanto, esta divisão relacionada ao lag dos métodos corresponderá a dois anos antes e dois anos depois do auge da crise de 2008 resultando 230 valores. Entretanto, para a série completa empregada no treinamento, será utilizada a primeira regra da Teoria das Ondas de Elliott, que corresponde ao fundo mais importante antes do auge da crise econômica que se inicia em 12/08/2002.

O Treinamento da primeira parte, antes do auge da crise de 2008, atenderá a 25% dos valores obtidos de um total de 460 valores do volume de 12/08/2002 até 30/05/2011; a Validação Cruzada corresponderá a 12,5% e a Verificação atenderá aos 12,5% do total. Na segunda parte, dois anos depois da crise de 2008, o Treinamento / Aprendizado atenderá a 25% dos valores obtidos do total restante, a Validação Cruzada corresponderá a 12,5% do total e a Verificação corresponderá aos 12,5% restantes de valores de entrada. Então, como pode ser observado nos parágrafos anteriores e no subcapítulo 2.2, a série utilizada compreenderá ao ciclo inteiro da crise de 2008. A Figura 15 (a) demonstra esta divisão de dados para a rede neural *Perceptron* de múltiplas camadas, e a Figura 15 (b), para a rede neural **Kohonen**.

⁵ Lag em séries temporais são operadores que defasam a variável, ou seja, tem relação ao tempo em que o método utilizará para fazer a sua previsão (AIUBE, 2007).

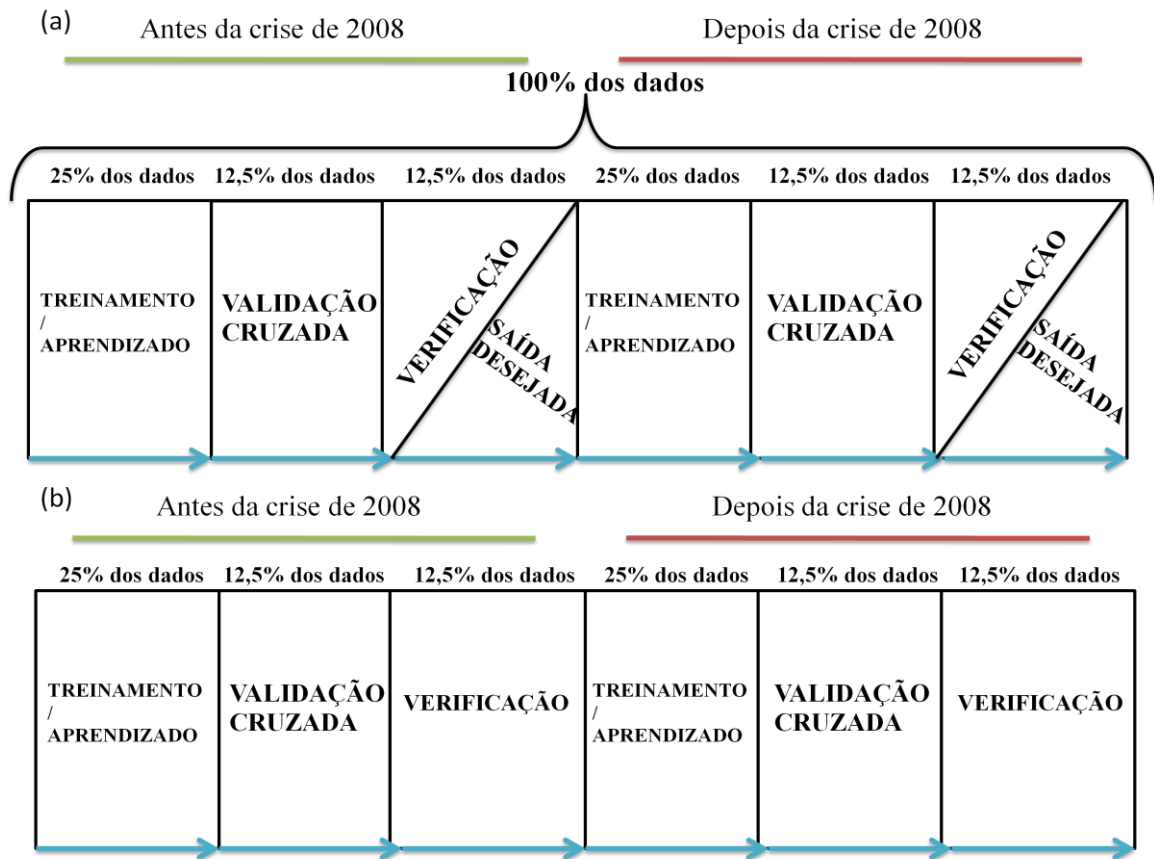


Figura 15 - Etapas da Manipulação dos Dados

Fonte: Organizada pelo autor

Como mencionado anteriormente, para que os valores sejam igualmente abordados no treinamento das duas redes neurais, necessitam serem transformados em valores correspondentes de entrada entre 0 e 1 (VALENÇA, 2009; SCHALKOFF, 2001). Com base nas afirmações acima, os valores recebidos no neurônio e nas entradas dos métodos serão normalizados segundo a Expressão (3.1) de normalização linear. Para que não haja uma diferenciação das entradas das redes neurais utilizadas, todas as saídas também serão normalizadas com a normalização linear, atribuindo uma melhor previsão para as ferramentas utilizadas (LUDWIG Jr. e MONTGOMERY, 2007).

$$x_{Norm} = \frac{x_n - x_{Mínimo}}{x_{Máximo} - x_{Mínimo}} \quad (3.1)$$

onde x_n representa a entrada; $x_{Máximo}$ representa o maior número da amostra de treinamento; $x_{Mínimo}$ representa o menor número da amostra de treinamento; e x_{Norm} representa o resultado. Quando utilizada a Expressão linear, predispõe-se uma confirmação de uma melhor

modelagem da série no treinamento, pois os dados de séries temporais não são intuitivos de modo que há intervalos muito distintos de variação, prejudicando a generalização da mesma.

Outra característica geralmente encontrada nas séries temporais é a raiz não unitária, que pode causar problemas na indução estatística da série, se não for ajustada corretamente. A série não estacionária de raiz da equação característica do processo diferente do valor um pressupõe que a média e a variância não devem ser constantes ao longo do tempo.

Então, para obter uma média e variância da série com raiz da equação com valor um, será aplicada uma equação para estimar $Y_z^j(x_n)$ (variáveis dependentes) com os valores defasados de x_{n-1} menos x_n (variáveis dependentes), conforme a Expressão (3.2) a seguir (FARIAS et al., 2010), onde α_n é a segunda constante; x_n e o valor atual da entrada; e x_{n-1} é o valor anterior:

$$Y_z^j(x_n) = \sum_{n=1}^p \alpha_n x_{n-1} - x_n \quad (3.2)$$

3.2 Métodos de Pré-treinamento Rede Neural MLP

Basicamente, todas as redes neurais são compostas por neurônio, camadas e sinapses, mas as grandes diferenças entre elas residem nas suas formas de treinamento / aprendizado e correção de sinapses. A escolha do número de neurônios de entrada (variáveis dependentes) e das camadas escondidas de uma rede neural depende da sua funcionalidade, propósito e desempenho.

Contudo, a escolha dos parâmetros modificáveis nas fases de funcionalidade e propósito de uma rede neural não é uma tarefa simples, pois é um processo de decisão difícil e requer um alto custo computacional, já que uma rede com valores errados reduz o aprendizado, de modo que não se possa ter uma garantia de previsibilidade (KATAGIRI et al., 2012). Durante o processo de desempenho, o objetivo é minimizar os erros entre o valor do resultado desejado $d_j(n)$ e o valor do resultado calculado pela rede neural $Y_z^j(n)$, conseqüentemente melhorando o ajuste das *sinapses* w_{ij} , e dos parâmetros modificáveis (AL-JUMAILY e CHEN, 2012).

Portanto, para cumprir o objetivo de identificação de crises econômicas, a rede neural *Perceptron* de múltiplas camadas será treinada com três neurônios de entrada que utilizarão três variáveis dependentes: o retorno do volume semanal do Ibovespa; o tempo; e a constante, como pode ser visto na Figura 16, a seguir. A utilização do retorno do volume de fechamento

do pregão semanal do Ibovespa corresponde à teoria das ondas, pois, para Elliott, o comportamento psicológico das massas está intrinsecamente ligado ao preço e ao volume dos ativos (NORONHA, 2009). O segundo neurônio da primeira camada da rede neural receberá o tempo correspondente ao número de semanas relativo aos mesmos valores do primeiro neurônio de todo o treinamento. A terceira variável independente da rede MLP será uma constante, com a finalidade de não originar um valor nulo para a função de transferência, tal que $\emptyset = \{c | c \neq 0\}$, pois valores nulos na fase de propagação progressiva do sinal (Fase Forward) Subcapítulo 2.3.4.2.1, ocasionam um não aprendizado para a rede neural.

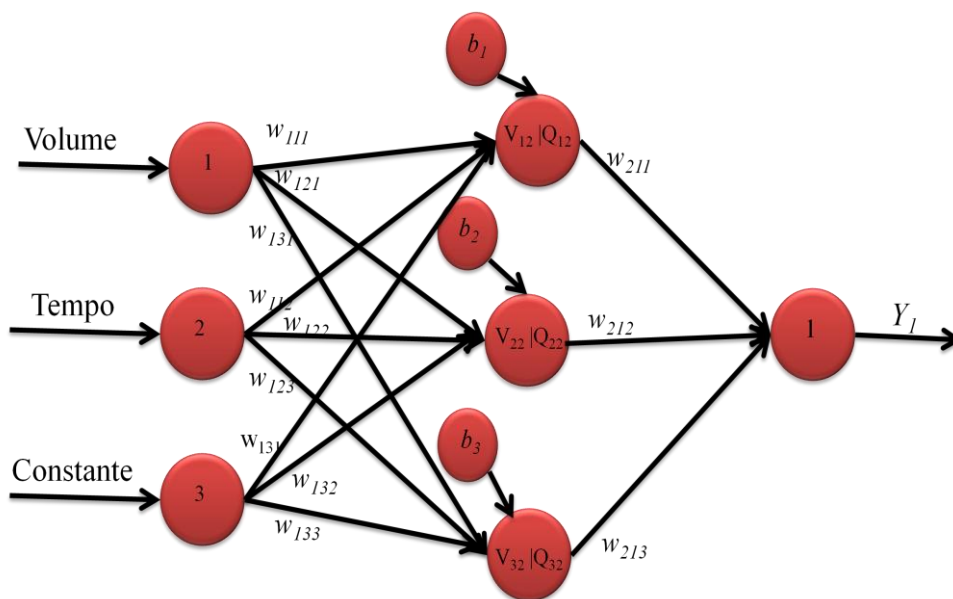


Figura 16 - Entradas da Rede Neural *Perceptron* de Múltiplas Camadas
Fonte: Organizada pelo autor

3.3 Métodos de Pré-treinamento Rede Neural Kohonen

Embora apresente as mesmas dificuldades das redes neurais MLP na escolha das variáveis dependentes, as redes **Kohonen** proporcionam um processo de decisão ainda mais difícil, pois são redes que buscam automaticamente um padrão aleatório nas entradas. Portanto, além de definir as entradas para a rede **Kohonen**, será necessário explicitar um padrão adotado que possa estar de acordo com as funcionalidades e os propósitos estabelecidos anteriormente, tornando possível a previsibilidade do objetivo.

Assim, para explicitar o padrão adotado no treinamento da rede neural **Kohonen**, será utilizado mais um neurônio de entrada correspondente à saída desejada utilizada na rede neural MLP. Ao contrário das redes neurais MLPs, a rede neural **Kohonen** apresentará, no

treinamento, quatro neurônios de entrada que utilizarão quatro variáveis dependentes: o retorno do volume semanal do Ibovespa; o tempo; a constante; e a saída desejada, como pode ser visto na Figura 17, a seguir.

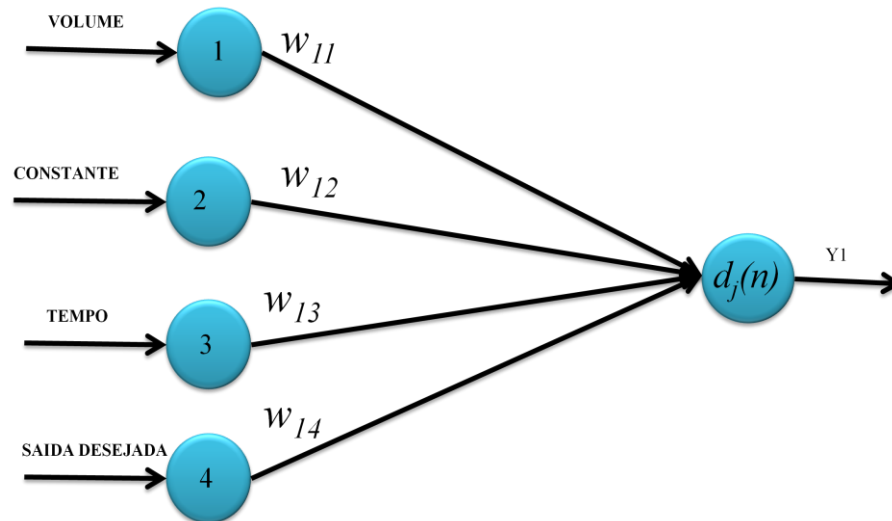


Figura 17 - Entradas da Rede Neural **Kohonen**

Fonte: Organizada pelo autor

Portanto, para a rede neural **Kohonen** estar de acordo com os parâmetros solicitados na revisão teórica, o primeiro neurônio de entrada de dados conterá o mesmo conjunto de valores do volume de fechamento semanal da rede neural MLP. No segundo neurônio de entrada da rede **Kohonen**, será auferido um vetor de valores correspondentes à constante da função, com a mesma finalidade citada anteriormente na rede neural MLP. Já no terceiro neurônio, a rede neural receberá o tempo correlativo ao número de semanas correspondente ao mesmo número de valores do volume de fechamento do pregão. Como referido anteriormente, a rede **Kohonen** terá um aprendizado não supervisionado, conseqüentemente, apresentará somente no treinamento um quarto neurônio contendo a saída desejada pelo supervisor.

3.4 Métodos de Treinamento

Após a escolha fundamentada no referencial teórico das variáveis dependentes no pré-treinamento, é necessário o desenvolvimento do padrão adotado pelo supervisor das redes neurais, baseado no objetivo da previsão. Neste caso, este padrão responsável pelo aprendizado denominado de saída desejada é baseado em um padrão da teoria das ondas de Elliott, subcapítulo 2.4, que tem como objetivo a identificação de crises econômicas.

Para a identificação de crises econômicas com base na teoria das ondas de Elliott, é utilizado o conhecimento do comportamento psicológico das massas, assim como os padrões relacionados a eles para compor a saída desejada das redes neurais. Para Elliott, é possível a previsão de padrões do comportamento psicológico das massas durante e após uma crise econômica. Portanto, para compor a saída desejada das redes neurais, será necessário delegar valores que explicitem o padrão da teoria das ondas de Elliott para cada uma das entradas que correspondem à série utilizada. Deste modo, no treinamento propriamente dito das duas redes neurais, será indispensável a escolha de pontos que simbolizem o final de uma onda e o começo da próxima, além de um ponto de venda que represente o término da quinta onda de impulso e o começo das ondas de correção.

Estes valores de entrada corresponderão ao vetor de saída desejada das duas redes neurais que será iniciado no começo dos valores de treinamento até o término do mesmo. Para compor os intervalos das ondas que não representam o final e/ou o começo da próxima onda, serão utilizados valores com coerência gradual, também baseados na teoria das ondas de Elliott, disposta no subcapítulo 2.4.

Como citado anteriormente, as duas redes neurais terão necessidade de um vetor de valores de entrada, responsável pelo aprendizado das redes, denominado de saída desejada. No caso da rede neural *Perceptron* de múltiplas camadas, este vetor de entrada será utilizado na comparação com a saída calculada no último neurônio como demonstrado no subcapítulo 2.3.4.2.2. Já na rede neural **Kohonen**, este vetor corresponderá a outro neurônio na camada de entrada, pois redes em que o aprendizado ocorre automaticamente baseiam-se somente em padrões aleatórios nas entradas para o seu treinamento.

3.4.1 Entradas: Saída Desejada

A saída desejada e uma entrada de dados ideal responsável pelo aprendizado da rede, ou seja, a entrada de dados em que a rede neural baseia-se para desenvolver a sua saída. Entretanto, para melhor entendimento de como se dará a construção do aprendizado para as redes neurais no vetor de saída desejada, este subcapítulo será dividido em três etapas que descreverão os padrões descritos no subcapítulo 2.4 da Teoria das Ondas de Elliott. Primeiramente, para o desenvolvimento do vetor da saída desejada das redes neurais, será feita uma classificação dos dados de entrada conforme os padrões encontrados na Teoria das Ondas de Elliott. A segunda etapa consistirá em qualificar cada entrada em uma nomenclatura própria que corresponderá a valores da saída desejada. E a terceira etapa

consistirá em identificar o que cada valor da saída desejada corresponde para cada entrada das redes neurais, ou seja, a cada valor do retorno do volume do Ibovespa.

3.4.1.1 Saída Desejada: 1ª Etapa

Na primeira etapa de desenvolvimento da saída desejada, foi feita uma divisão da entrada das redes neurais em duas partes, antes e depois do auge da crise. Para cada parte foram propostas mais duas divisões, antes do auge da crise, em cinco ondas de impulso, e depois do auge da crise, em três ondas de correção (ANEXO A). Para as duas classificações (divisões) dos dados de entrada, levam-se em conta todos os conceitos de desenvolvimento de ondas descrito no subcapítulo 2.5.

Para melhor entendimento da execução dos conceitos de identificação de ondas, será desenvolvida uma sequência de figuras que demonstrarão a aplicação dos mesmos no cenário real proposto. Na Figura 18, a seguir, pode-se observar a realização da terceira regra, referente à identificação da primeira onda, que geralmente se inicia em um fundo importante no mercado, neste caso, está circulado o fundo de 12/08/2002. Neste período também é iniciada a entrada do treinamento das redes neurais.

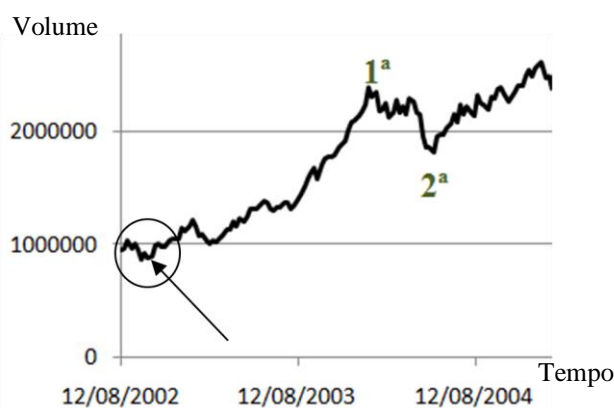


Figura 18 – Aplicação da terceira regra de Elliott

Fonte: Organizada pelo autor

Na Figura 19, pode-se observar a execução da quinta regra relacionada com a ocorrência de prováveis extensões em uma das ondas ímpares. No período utilizado para o treinamento das redes neurais, foi identificada uma extensão na quinta onda, que se iniciou no final da quarta onda, ocasionando um ziguezague até o término da quinta onda. Este ziguezague que ocorreu no período de formação da quinta onda pode ser denominado de

extensões da quinta onda, pois, ao invés de findá-la no ponto (I), o mercado ainda se estendeu para mais uma contagem de ondas de grau menor como demonstrado na Figura 19, a seguir.

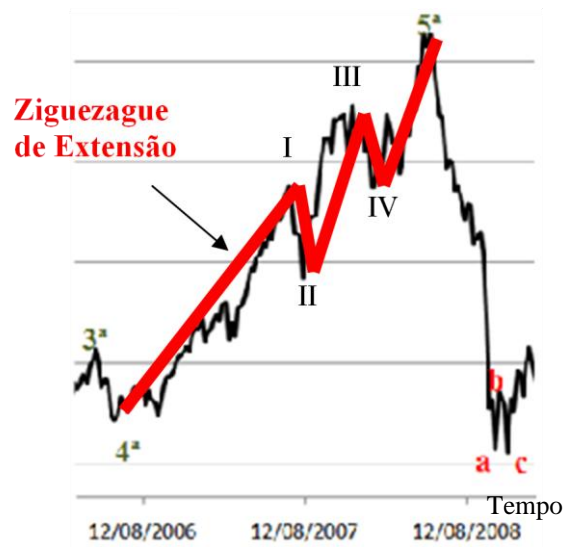


Figura 19 – Aplicação da quinta regra de Elliott

Fonte: Organizada pelo autor

A sexta etapa a ser desenvolvida foi referente à identificação da segunda onda, que não obteve a ocorrência demonstrada na regra descrita no subcapítulo 3.4.2, pois, ao invés de corrigir de 50 a 62% do tamanho da onda um, como ocorrido em 73% dos casos, a onda dois corrigiu de 32 a 50% da onda um, respeitando 12% das vezes do ocorrido. A Figura 20, a seguir, demonstra a relação entre a porcentagem de ocorrência com a porcentagem da correção da onda dois, onde podemos observar que a retração da onda dois ficou entre 32 a 50% da onda um, caracterizando uma onda pequena com pouca força de vencedores.

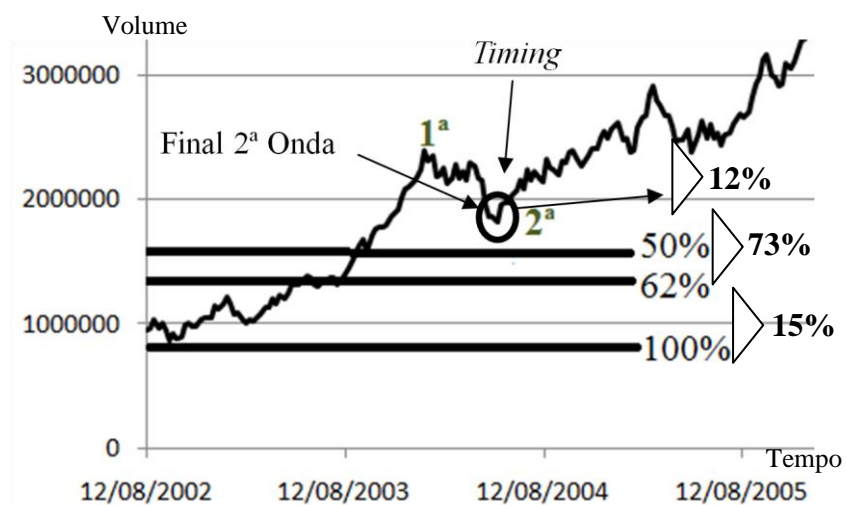


Figura 20 – Aplicação da sexta regra de Elliott

Fonte: Organizada pelo autor

Na Figura 21, a seguir, pode-se observar, em um cenário real, o desenvolvimento da sétima regra, que faz referência à identificação do final da terceira onda. No período utilizado no treinamento das redes neurais para identificar o término da terceira onda, foi utilizado o critério da razão do número de ouro. Assim, o tamanho da onda três deverá ter uma razão de 1,618 (número de ouro) até 1,75 vezes do tamanho da onda um. Na Figura 21, pode-se observar esta relação, onde o topo da onda três respeita o intervalo de 1,618 até 1,75 descrito na regra da teoria das ondas de Elliott.



Figura 21 – Contagem da onda quatro, oitava regra da teoria das ondas.

Fonte: Organizada pelos autores.

Ainda na Figura 21, pode-se observar que o fundo da onda quatro não ultrapassa o topo da onda um, respeitando a oitava regra da teoria das ondas de Elliott. Esta regra está definida com uma reta que corta a Figura 18 em duas partes, o que demonstra a validade na contagem de ondas até a onda quatro (NORONHA, 2009).

A nona e última regra descrita neste trabalho, no subcapítulo 3.4.2, faz alusão à identificação do final da quinta onda. Na Figura 22, pode-se analisar a relação desta regra no cenário real utilizado, onde, para a identificação do término da quinta onda, foram utilizadas duas retas: a primeira, com o critério de transpor pelo máximo de topos possíveis e segunda reta, pelo máximo de fundos possíveis entre a onda três e a onda cinco. Para finalizar a quinta onda, foi utilizado o ponto em que a cotação ultrapassou a reta de fundos, ou seja, quando a segunda reta foi rompida pela cotação, simbolizando o final da quinta onda. Para melhor reconhecimento do ponto adequado para o final da quinta onda, circulou-se no gráfico o rompimento da segunda reta, como demonstrado na Figura 22, a seguir.



Figura 22– Triângulo da nona regra da Teoria das Ondas de Elliott

Fonte: Organizada pelos autores.

Portanto, ao término da aplicação da nona regra, pode-se dar por finalizado o ciclo e o final da contagem das ondas de impulso (Mercado de Alta). Como se pode observar nos parágrafos anteriores, para a identificação das ondas no período utilizado para o treinamento das redes neurais, foram usadas todas as nove regras descritas no subcapítulo 3.4.2. Entretanto, como mencionado anteriormente, em outras circunstâncias estas regras poderiam variar de acordo com as variações encontradas em Noronha (2009), Chaves (2003), Elliott (2011) ou Debastiani (2008). Com a demonstração destes padrões para as redes neurais, será possível “ensiná-las” a operarem no mercado de ações, mediante uma crise econômica, classificada com o mesmo agente causador da crise de 2008 mencionado no subcapítulo 2.2.

3.4.1.2 Saída Desejada: 2ª Etapa

Após a contagem das ondas seguindo os padrões destacados na teoria das ondas de Elliott, será necessário qualificar cada entrada em uma nomenclatura própria que corresponda a valores da saída desejada que a rede neural compreenda. Para isso será utilizado o critério de que, se a saída da rede corresponder a um resultado próximo ao valor 1 ou o próprio 1, o prognóstico será de venda ou topo da crise, descrita na nona regra, subcapítulo 3.4.2.

Quando este resultado corresponder a valores distintos do valor 1, o prognóstico será de sustentação, ou seja, de construção da ondas um, dois, três e quatro. Estes pontos

proeminentes do valor 1 serão organizados com coerência gradual para intervalos de valores de 0,01618 para ondas ímpares e -0,01618 para ondas pares (número de ouro na escala de 1:100.000).

Para que a rede neural entenda a diferença de cada onda, será utilizada a classificação por escala, ou seja, a identificação de onda será dada por categorias distintas graduais. Então, para o primeiro par de ondas, foram utilizados intervalos de valores correspondentes a 0,00001618, para a primeira onda, os intervalos de valores de -0,00001618 para a segunda onda, assim por diante relacionando o intervalo de valores para cada par de onda, como mostra a Tabela 6, a seguir, para melhor entendimento, com um fragmento aleatório fictício da entrada. Entretanto, estes valores da saída desejada somente serão testados no treinamento das redes, podendo ser modificados no decorrer da calibração.

Tabela 6- Entradas Saída Desejada

ID	Onda	Saída Desejada	Classificação
1	3	0,0001618	Sustentação
2	3	0,0001618	Sustentação
3	4	-0,0001618	Sustentação
4	4	-0,0001618	Sustentação
5	4	-0,0001618	Sustentação
6	4	-0,0001618	Sustentação
7	5	0,001618	Sustentação
8	5	0,001618	Venda
9	1	0,000001618	Sustentação
10	1	0,000001618	Sustentação
11	1	0,000001618	Sustentação
12	2	-0,000001618	Sustentação

Fonte: Organizada pelos autores.

Na Tabela 6, semelhantemente às Tabela 4 e 5, pode-se observar três colunas de entradas id, onda e saída desejada e uma coluna de resultado, classificação. Entretanto, estes valores deverão ser codificados de uma forma que as redes neurais entendam o padrão requerido. Para isso, deverão ser transformados em três colunas para as redes neurais **Perceptron** de múltiplas camadas e quatro colunas para as redes neurais **Kohonen**.

3.4.1.3 Entrada das Redes Neurais: 3ª Etapa

Como se expôs anteriormente, as redes neurais serão treinadas a partir de uma gama de valores de entrada, contendo e correspondendo aos conceitos da Teoria das Ondas de Elliott, subcapítulo 3.4.2. Porém, para que a rede neural compreenda os valores de entrada, cada valor foi decodificado na sua coluna de entrada correspondente, ou seja, ao invés de utilizar os mesmos valores da Tabela 6, estes valores foram decodificados para as redes neurais.

Primeiramente, para a modificação dos valores das colunas de entrada, foi utilizando o processo de diferenciação dos valores Expressão (3.2) e, em um segundo estágio, a Expressão (3.1) de normalização. Para as colunas de entrada do retorno do volume de fechamento semanal do Ibovespa e a do tempo, foram utilizadas as duas equações. Porém, para a saída desejada, constante e o tempo não foi utilizada nenhuma equação de modificação, pois os valores já estavam em correspondência com os requisitos de valores de entrada das redes neurais, como mostra a Tabela 7, a seguir.

Tabela 7- Entradas Decodificadas para as Redes Neurais

Volume	Constante	Tempo	Saída desejada
0,438172	1	0,001	0,000001618
-0,36918	1	0,002	0,000003236
0,072581	1	0,003	0,000004854
0,270609	1	0,004	0,000006472
-0,25717	1	0,005	0,00000809
-0,04749	1	0,006	0,000009708
-0,08513	1	0,007	0,000011326
0,250896	1	0,008	0,000012944
-0,15143	1	0,009	0,000014562
-0,68369	1	0,010	0,00001618

Fonte: Organizada pelos autores.

3.4.2 Treinamento / Aprendizado Rede Neural *Perceptron*

Após a decodificação dos dados no pré-treinamento/aprendizado, foi utilizada a modificação de quatro parâmetros modificáveis buscando obter um erro aceitável, mencionado no subcapítulo 2.3.6.2 para o treinamento/aprendizado da rede neural *Perceptron*: o número de camadas escondidas; a taxa de aprendizado; a taxa do erro; e os

ciclos. No entanto, o valor final destes parâmetros modificáveis somente será obtido ao término do treinamento, pois todos os parâmetros mutáveis de uma rede neural serão testados a partir do método de tentativa e erro baseado em Sahoo et al., (2012).

O número de camadas escondidas diz respeito à topologia da rede neural que demonstra quantos neurônios cada camada terá. A taxa de aprendizado é um parâmetro dado pelo supervisor que se relaciona com o valor de aprendizado do ciclo inteiro de dados de uma rede neural, ou seja, a rede neural aprenderá no máximo o valor da taxa de aprendizado no ciclo inteiro. Entretanto, uma taxa de aprendizado muito baixa torna lento o treinamento, e uma taxa de aprendizado muito alta impede a tendência do processo de aprendizado de uma rede neural.

O erro é uma taxa pré-definida responsável pela generalização da rede que encerra o treinamento após o valor atingido. Já o ciclo é o número de vezes em que o conjunto de dados do treinamento é apresentado à rede neural. Assim como a taxa do erro e o ciclo, são considerados critérios de parada da rede neural. Porém, existem vários outros critérios já mencionados. Todos estes parâmetros deverão estar em acordo com os erros já citados no subcapítulo 2.3.6.2, para que a rede neural esteja treinada corretamente, isto é, a modificação adequada destes quatro parâmetros leva a rede neural a obter um erro aceitável baseado no método de tentativa e erro.

3.4.3 Métodos da Validação Cruzada e Verificação

Após o término do treinamento/aprendizado das redes neurais com a modificação dos parâmetros na fase anterior, será necessário o desenvolvimento da análise da validação cruzada, que proporcionará uma melhor aceitação dos resultados. Esta análise dos dados da validação consistirá na utilização de uma equação em dois estágios distintos como um critério de parada aceitável: A Equação (2.7) do erro médio quadrático do treinamento/aprendizado; e a Equação (2.7) do erro médio quadrático da validação cruzada. Para os dados de entrada utilizados nesta fase, são empregados valores de entradas distintas das do treinamento, garantindo uma melhor generalização dos resultados.

Dessa forma, posteriormente à análise da validação cruzada, poder-se-á concluir se as sinapses e bias das redes neurais corresponderão a uma relação válida entre a média de erro nas duas etapas anteriores. Entretanto, esta conclusão somente será aceitável se os padrões propostos no treinamento/aprendizado e na validação cruzada continuarem os mesmos, ou seja, a reta do erro da validação cruzada deverá cruzar a reta do erro do

treinamento/aprendizado (BARBOSA et al., 2005), como demonstrado no subcapítulo 2.3.6.2. Conclui-se, portanto, que esta modelagem dos parâmetros modificáveis das redes neurais na etapa da validação cruzada deverá proporcionar um resultado tão generalizável quanto possível, ou seja, a modelagem dos parâmetros servirá para entradas que não estavam presentes no treinamento.

A última etapa do treinamento a ser concluída será a verificação. Nesta fase, poderá ser constatado que os métodos não perderão sua precisão na troca de entradas, pois o erro médio quadrado, Equação (2.7), terá que permanecer entre os valores aceitáveis, conforme o subcapítulo 2.3.6.2. Nesta etapa, haverá a comparação propriamente dita da previsão das duas redes neurais, utilizando o critério do Erro Médio Quadrático e do R^2 ajustado. No primeiro critério, busca-se selecionar a rede neural que obtiver o menor Erro Médio Quadrático na verificação, pois, quanto mais baixo o erro for, melhor será seu poder computacional e mais bem treinada estará a rede neural. O segundo critério a ser analisado será o R^2 ajustado, onde se busca selecionar a rede neural que obtiver o resultado mais próximo do valor 1, demonstrando que a rede neural vencedora terá maior poder explicativo das variáveis dependentes com as variáveis independentes.

3.5 Testes utilizados

Neste último subcapítulo do método do trabalho, serão apresentados os testes a serem desenvolvidos para sanar ou minimizar ao máximo os problemas das séries temporais na utilização das redes neurais. Estes testes serão aplicados após a verificação cruzada, a rede neural que obtiver o melhor desempenho no conjunto do treinamento/aprendizado, validação cruzada e verificação.

O tratamento dos dados em relação aos testes utilizados é um passo extremamente negligenciado por muitos pesquisadores das áreas exatas mesmo sendo considerado como uma das partes mais importantes na validação dos modelos heurísticos. Alguns dos grandes problemas de se utilizar o retorno do volume das negociações é a consideração de que as séries temporais, provenientes do mercado de renda variável em quase a sua totalidade, não são estacionárias, com autocorrelação e heterocedasticidade nos resíduos (ALVES, 2007; BARBA, 2011). Mas, para Alves (2007) e Cargnelutti et al., (2009), as séries temporais podem proporcionar uma predição confiável se cumpridos seus requisitos corretamente.

A análise dos dados será realizada por meio do software R versão 2.13.2, utilizando os pacotes necessários dos testes utilizados. A análise dos resultados após a previsão será

auferida por meio de três testes: Garson (1991) e Goh (1995); Breusch-Pagan (BREUSCH e PAGAN, 1979); White (WHITE, 1980); e Durbin e Watson (1950).

O teste de Garson (1991) e Goh (1995), subcapítulo 2.3.8, será desenvolvido para eliminar o problema da “caixa preta” das redes neurais. Este teste influencia na escolha das variáveis de entrada adotadas a partir das variáveis brutas, eliminando as variáveis e neurônio que menos se refletem no resultado.

Empregando-se mais dois cálculos no algoritmo *backpropagatio* após a correção do erro dos neurônios, a equação da sensibilidade descrita por Garson, em 1991 (3.3 a), e a equação da contribuição percentual delineada na Expressão 3.3 b, a seguir, pode-se descobrir a influência de cada variável dependente em relação às variáveis independentes e, assim, eliminar as variáveis que menos contribuem para o resultado (VALENÇA e LUDERMIR, 2007).

$$Sen_{ji} = \sum_{k=1}^{n_{hid}} w_{kj} f'(net_k) \cdot \sum_{i=1}^{n_{out}} w_{ik} f'(net_i) \cdot e_i ; \quad (3.3 a)$$

$$Cont_j(\%) = \frac{\sum_{i=1}^N Sen_{ji}^2}{\sum_{j=1}^{N_{inp}} \sum_{i=1}^N Sen_{ji}^2} ; \quad (3.3 b)$$

onde: Sen_{jL} é a sensibilidade para cada variável j ($j = 1, \dots, N_{inp}$) de entrada com relação à saída para um dado exemplo i ($i = 1, \dots, N_{inp}$) é o número de neurônios na camada de entrada; N é o número de exemplos utilizados no treinamento; N_{out} é o número de neurônio na camada de saída, w_{ik} são os pesos sinápticos que conectam a camada de saída a camada escondida; $f'(net_i)$ é a derivada da função de ativação dos neurônios da camada de saída; e_i é o erro; N_{hid} é a quantidade de neurônios na camada escondida onde ($k=1, \dots, N_{hid}$); w_{kj} são os pesos sinápticos que conectam a camada escondida à camada de entrada; e $f'(net_k)$ é a derivada da função de ativação dos neurônios da camada escondida.

Para verificar a heterocedasticidade do resíduo da amostra, será desenvolvido dois teste já citados anteriormente, o teste Breusch-Pagan e o teste White (1980), por meio dos quais serão verificados os resíduos com a variância não constante para todas as observações, ou seja, uma rejeição de H_0 com os resíduos homocedásticos. O critério para solução do problema será a exclusão de variáveis problemáticas dos construtos primitivos.

O último teste a ser analisado será o de Durbin-Watson, que avaliará a autocorrelação das variáveis e construtos, ou seja, em um modelo, pressupõe-se que o termo de erro de

qualquer variável não é influenciado pelo termo de erro de suas conseqüentes e nem o influencia. Para uma rejeição de H_0 , os resíduos deverão ser não autocorrelacionados, e o critério para solução do problema será a exclusão de variáveis problemáticas dos construtos iniciais (ALVES, 2007). Em face do que já se disse, a rede neural será considerada treinada, validada e testada até esse exato momento, o que viabilizará as apresentações posteriores dos resultados.

4 ANÁLISE E DISCUSSÕES DOS RESULTADOS

Neste capítulo serão apresentados os resultados empíricos encontrados para os dados utilizados neste estudo. Basicamente, este capítulo será dividido em três etapas: a primeira corresponde a apresentação dos resultados do treinamento / aprendizado, validação cruzada das duas redes neurais; em seguida será examinada a etapa mais conclusiva do trabalho, a verificação dos métodos, onde haverá uma comparação da previsão das redes neurais MLP e **Kohonen**; e, enfim, na última etapa será demonstrado a análise dos resultados dos testes utilizados na rede neural que obteve o conjunto treinamento e verificação com os menores erros.

No caso de métodos heurísticos como as redes neurais, a busca de um percentual de erro aceitável descrito no subcapítulo 2.3.6.2 na modelagem dos parâmetros no treinamento/aprendizado e validação, tem a finalidade de proporcionar um modelo de previsão adequada com o treinamento satisfatório. Tais modelos podem ser classificados como uma classe de modelos não-lineares, que permitem ao supervisor o ajuste de um grande número de parâmetros modificáveis e diferentes configurações (BRESSAN, 2004). Essas diversas configurações podem ser decodificadas como formas operacionais alternativas para um ajuste adequado da rede em termos da topologia escolhida, da taxa de aprendizado, da taxa do erro, e do número de ciclos (FERNANDES, 1995).

A construção de um modelo de redes neurais envolve desde a modelagem adequada da rede até as transformações utilizadas para transmitir os dados a ela e aos métodos utilizados para interpretar os resultados obtidos. Esses três aspectos (modelagem, transformações e interpretação) são fundamentais na utilização das redes neurais para previsão (BRESSAN, 2004). Então, com base nas afirmações dos parágrafos anteriores este capítulo e subcapítulos seguintes conterão uma análise interpretativa dos resultados obtidos na modelagem e transformações dos parâmetros modificáveis das duas redes neurais.

4.1 Rede neural MLP Treinamento e Validação Cruzada

A primeira rede a ser analisada será a rede neural MLP com o aprendizado supervisionado descrito no subcapítulo 2.3.4.2. Como já foi mencionado anteriormente para transformar a série utilizada pela rede neural MLP, em uma série com raiz unitária. Utilizou-se a Equação (3.2) de diferenciação da série original com o retorno da mesma em t_{n-1} . A Figura 23, a seguir, demonstra graficamente o índice do retorno utilizado.

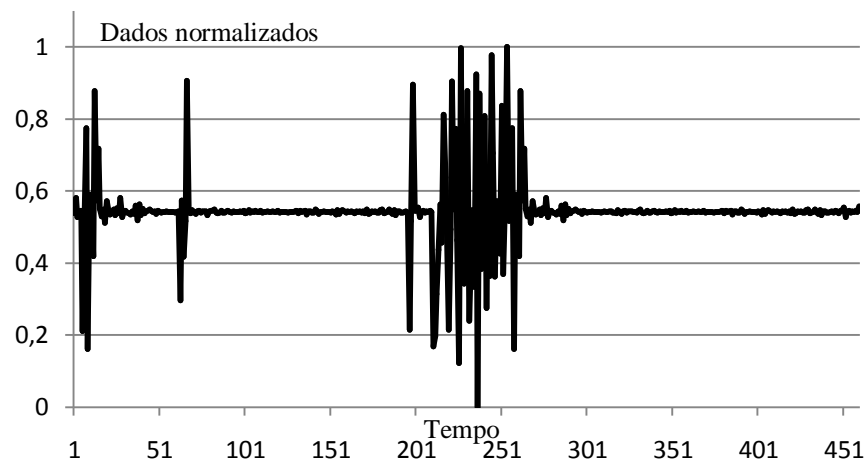


Figura 23 – Série retorno do volume semanal normalizada com raiz unitária

Fonte: Organizada pelos autores.

Na Figura 23, pode-se concluir que os retornos dos índices são estacionários, pois se desenvolveram ao longo do tempo, ao redor de um valor estável, apresentam uma forma de equilíbrio constante, corroborando com Farias (2008), que utilizou esta mesma técnica do retorno para detectar estacionalidade entre as entradas. Após o teste gráfico de estacionalidade dos dados, é feita a modificação dos parâmetros de uma rede neural leva-a à aquisição do conhecimento ou aprendizado do determinado objeto. Esta aquisição gradual do conhecimento ocorreu no seguimento de cinco grandes etapas da modelagem de uma rede neural: a modificação da taxa de aprendizado α ; a escolha da topologia da rede neural $TP_{p,e,s}$; a escolha do erro final $e_j(n)$; e a escolha do ciclo utilizado.

Outro parâmetro pode ser levado em consideração, à escolha do intervalo da saída desejada. Entretanto, somente o valor deste parâmetro não será testado, pois será utilizado o número de ouro 1,618 em escala de (1:1.000.000) descrito no Subcapítulo 3.4.1.2, que diz respeito ao intervalo de valores correspondentes a cada valor na teoria das ondas de Elliot.

Então, o primeiro parâmetro a ser levado em consideração no treinamento será a taxa de aprendizagem por interação que foi escolhida a partir dos testes feitos na etapa do treinamento da rede neural. Neste processo ocorreu uma comparação entre os valores escolhidos com os resultados dos erros obtidos. A Tabela 8 a seguir, demonstra os critérios utilizados que foram o menor erro médio global e o menor erro médio quadrático. Além das taxas dispostas na Tabela 8, outras ainda foram testadas.

Tabela 8 - Taxa de Aprendizado

	EMG no Ciclo 100	EMQ no Ciclo 100
Taxa 2	0,00925	0,00516
Taxa 1	0,01683	0,00877
Taxa 0,8	0,00885	0,00470
Taxa 0,5	0,10392	0,01604

Fonte: Organizada pelos autores

Como pode ser visto na Tabela 8, com os testes realizados concluiu-se que a taxa de aprendizado por interação 0.8, teve o melhor custo benefício entre as testadas, com o menor erro médio global e menor erro médio quadrático no ciclo 100. O segundo critério a ser analisado para o treinamento da rede neural foi à modificação da topologia, ou seja, a quantidade de neurônios na camada de entrada, na camada escondida e na camada de saída.

Como teste realizou-se a comparação das varias modificações neste parâmetro. No entanto a topologia que mais se destacou foi a 3-8-1, pois foi à melhor resposta no treinamento dentre as sete diferentes topologias testadas, variando entre 2 a 10 neurônios na camada oculta. A melhor situação correspondeu a três neurônios na camada de entrada, oito neurônios na camada oculta e um neurônio na camada de saída.

Nessa etapa também foi observado quantos ciclos seria necessários para que a predição ocorresse em um percentual de erro médio quadrático aceitável, como mostra a Tabela 9, a seguir, com os resultados mais relevantes do erro médio quadrático na saída da rede neural. Como podemos observar na Tabela 9 em negrito, o ciclo 100 foi o que mais se destacou segundo o critério de menor erro médio quadrático com a topologia 3-8-1.

Tabela 9 - Tabela da Topologia da MLP

	3-4-1	3-6-1	3-8-1	3-10-1
Ciclo 100	0,00786	0,00532	0,00470	0,06913
Ciclo 500	0,10856	0,06283	0,07899	0,08581
Ciclo 1000	0,01148	0,09796	0,06882	0,10589

Fonte: Organizada pelos autores.

Na Tabela 9, pode se observar que um ótimo resultado do erro médio quadrático foi obtido no ciclo 100 com seis neurônios na camada oculta. Porém, a topologia que mais se destacou segundo o critério do menor erro médio quadrático foi à topologia 3-8-1 com 100

ciclos em negrito. Como mencionado anteriormente, além das topologias e ciclos dispostos na Tabela 9, outras ainda foram testadas.

O ultimo critério para o treinamento de uma rede neural foi o valor do erro desejado por interação. Este erro é um valor escolhido pelo supervisor que compreende o máximo de erro suportado pela rede neural em cada interação, ou seja, o produto da Equação (2.4) de cada interação da rede neural não deverá ser maior que o valor escolhido ocasionando uma parada no treinamento.

Entretanto, esse valor não deverá ser muito baixo e nem muito alto, pois a rede não ira aprender o necessário no primeiro caso e no segundo estará super-treinada para a entrada correspondente. Então, para a escolha do valor do erro desejado por interação foi utilizado três critério, o menor erro médio quadrático, o menor erro médio global, e o menor erro padrão de previsão, sem danificar os resultados obtidos anteriormente. A Tabela 10 a seguir, demonstra os resultados dessa comparação entre as escolhas do valor de erro.

Tabela 10 - Erro Desejado

	EMQ	EMG	EP
Erro Desejado 0,70	0,00795	0,01020	0,08919
Erro Desejado 0,50	0,00470	0,00885	0,06861
Erro Desejado 0,30	0,00588	0,00970	0,07673
Erro Desejado 0,09	0,00612	0,00980	0,07825

EMQ = Erro Medio Quadratico; EMG = Erro Medio Global; EP = Erro Padrão de Previsão.

Fonte: Organizada pelos autores.

Na Tabela 10, concluiu-se que o valor do erro desejado 0,50, teve o melhor custo benefício entre os valores comparados, com a melhor aproximação do erro desejado e o menor erro propiciado pelo esforço da rede neural para alcançá-lo. Levando-se em consideração o treinamento inteiro, a rede neural com 100 ciclos topologia 3-8-1, taxa de aprendizagem 0,8 e erro desejado de 0,5 foi considerada a mais satisfatória no conjunto do treinamento.

No ANEXO B, podemos observar dois gráficos (a) e (b). No gráfico (a), para melhor visualização, está exposta a cotação do Ibovespa, classificada em ondas, correspondente ao tempo do resultado do treinamento da rede neural MLP. No gráfico (b), podemos observar a comparação do resultado do treinamento dos quatro melhores erros dispostos na Tabela 10, onde se pode concluir, que foi possível a identificação das mesmas cinco ondas dispostas no gráfico (a) no ANEXO B, seguindo o treinamento disposto no método do trabalho.

Para a validação cruzada da rede neural foi utilizado um gráfico contendo o erro calculado pela rede neural no treinamento em comparação com o erro calculado pela rede neural na validação cruzada. Esta comparação, como mencionado anteriormente no subcapítulo 2.3.6.2, servira para localizar o ponto ótimo de parada do treinamento dentre todos os testados, ou seja, será demonstrado na Figura 24, exatamente o ciclo em que a rede neural estará com melhor treinamento pra a série de dados utilizada.

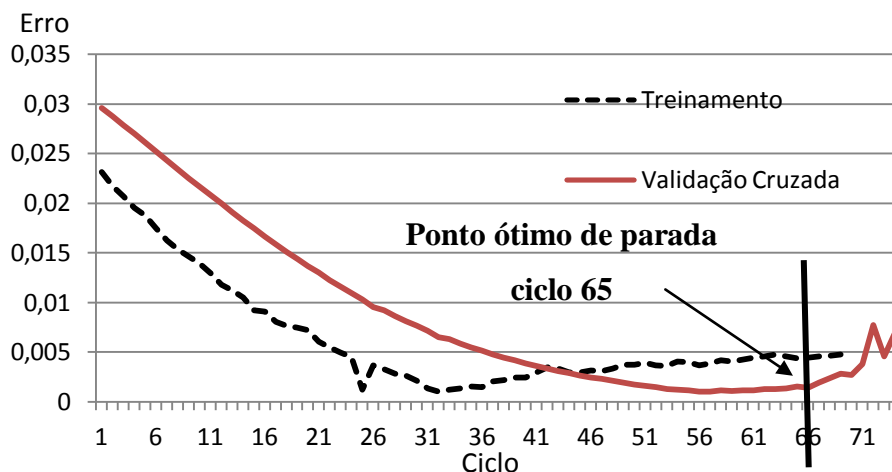


Figura 24 – Treinamento VS Validação Cruzada

Fonte: Organizada pelos autores.

Então, como demonstrado na Figura 24, concluiu-se que a rede neural com 65 ciclos topologia 3-8-1, taxa de aprendizagem 0,8, e erro desejado de 0,5, foi considerada a configuração mais satisfatória no conjunto do treinamento e validação cruzada. Como decorrência disso, a configuração mencionada anteriormente será utilizada na validação da rede neural, onde haverá uma comparação entre as duas redes neurais desenvolvidas.

4.2 Rede neural Kohonen Treinamento e Validação Cruzada

O segundo resultado do treinamento a ser analisado será o da rede neural **Kohonen** com o aprendizado não supervisionado descrito no subcapítulo 2.3.5.5. Como referido previamente, a modificação dos parâmetros de uma rede neural leva-a a aquisição do conhecimento ou aprendizado do determinado objeto. Entretanto, para redes com aprendizado não supervisionado o supervisor não é responsável pela modificação dos parâmetros citados no subcapítulo 4.1, porém é responsável pelos valores iniciais destes parâmetros. Então, a aquisição do conhecimento de uma rede neural **Kohonen** ocorre no seguimento de sete grandes etapas da modelagem automática: a taxa de aprendizado inicial ∞ ; a topologia inicial

da rede neural $TP_{p,s}$; a escolha do erro inicial $e_i(n)$; o ciclo inicial utilizado; o valor inicial da vizinhança topológica h_{ik} ; e o tempo final de execução da rede t_f .

Convém ainda lembrar, que os dados utilizados na rede neural Kohonen foram os mesmo utilizados na rede neural MLP, assim como os valores testados dos parâmetros finais da rede neural MLP, taxa de aprendizado, topologia, erro e ciclo utilizado. Porém, somente os parâmetros específicos da rede neural **Kohonen**, tempo final de execução que foi fixado em 60 minutos e o valor da vizinhança topológica inicial em 0,2, que foram dados no começo do treinamento/aprendizado.

Portanto, a rede neural **Kohonen** utilizou no seu treinamento/aprendizado e validação cruzada os valores iniciais dos parâmetros como ciclo 65, topologia 3-1, taxa de aprendizagem 0,8, erro desejado de 0,5, parâmetro de tempo final de execução 60 minutos e valor da vizinhança topológica inicial em 0,2. Como podemos observar no parágrafo anterior o parâmetro topologia foi reduzido a somente camada de entrada e camada de saída, pois este tipo de rede neural não compreende camadas ocultas como as redes MLP.

A Tabela 11 a seguir, demonstra os resultados obtidos no final da validação cruzada da rede neural **Kohonen** que por sua vez ajustou ainda mais os parâmetros iniciais utilizando ciclo 65, topologia 3-1, taxa de aprendizagem 0,8, erro desejado de 0,52, parâmetro de tempo final de execução $t_1 = 60$ minutos e valor da vizinhança topológica final em 0,46. Resta afirmar que os erros obtidos na validação cruzada da rede neural **Kohonen** são superiores aos erros do treinamento/aprendizado da rede neural MLP. Porém, a afirmação de que a previsão da rede neural **Kohonen** é superior a previsão da rede neural MLP ainda não é conclusiva. Nesse caso, esta afirmação somente será corroborada na verificação das duas redes, onde novas entradas serão dispostas a fim de sumarizar aos resultados já pré-estabelecidos.

Tabela 11 - Validação Cruzada Rede Neural **Kohonen**

	Rede Kohonen
EMG	0,00588
EMQ	0,00328
EP	0,05728
R ² Ajustado	0,03298

Fonte: Organizada pelos autores.

Na Tabela 11, podemos observar nos três primeiros valores, equações referentes ao erro da previsão e no ultimo valor uma equação relativa à avaliação de resultados. Então, deve-se acrescentar, enfim, que a rede neural **Kohonen** apresenta resultados acima dos

definidos no subcapítulo 2.3.6.2, demonstrando uma modelagem muito satisfatória. Entretanto, apesar dos ótimos resultados da rede neural **Kohonen** na validação cruzada, é possível afirmar que a um grande problema na utilização de dados temporais em redes não supervisionadas. Este problema refere-se à procura de um padrão parcialmente aleatório por essas redes, sendo esse, muito difícil de ser descoberto pelo supervisor. Então, a partir desta constatação não há como saber que padrões a rede neural **Kohonen** utilizou para chegar a esses valores de erro, deixando seus resultados ainda mais imprevisíveis na verificação das duas redes neurais.

4.3 Verificação

Como referido anteriormente, após a verificação cruzada das duas redes neurais será necessário validá-las. Entretanto, para isso em primeiro lugar será imprescindível apresentar dados distintos do treinamento/aprendizado e verificação cruzada para as redes neurais testados nos subcapítulos anteriores. Em segundo lugar, será feita uma análise da previsão do retorno do volume das duas redes neurais, buscando prever uma crise econômica. Em seguida, será apresentado o desenvolvimento de uma comparação entre os resultados obtidos seguindo alguns critérios, a fim de, demonstrar qual das redes neurais tem mais pré-disponibilidade na tomada de decisão em relação ao mercado acionário em épocas de crises financeiras. E, finalmente, algumas discussões dos resultados obtidos na verificação.

A verificação de uma rede neural é uma das etapas mais determinantes do treinamento, pois, avalia o desempenho e valida o modelo treinado. Assim como em todas as etapas do treinamento, na verificação, o conjunto de dados de entrada utilizado pelas redes neurais deverá ser distinto. Em consequência disto, foram utilizados para a previsão, os 115 valores restantes do retorno do volume de fechamento semanal do Ibovespa de 02/03/2009 até 30/05/2011. Na Figura 25, a seguir, para melhor entendimento do período utilizado, podemos observar um gráfico com a cotação do Ibovespa no mesmo período empregado como variável de entrada na verificação das redes neurais.

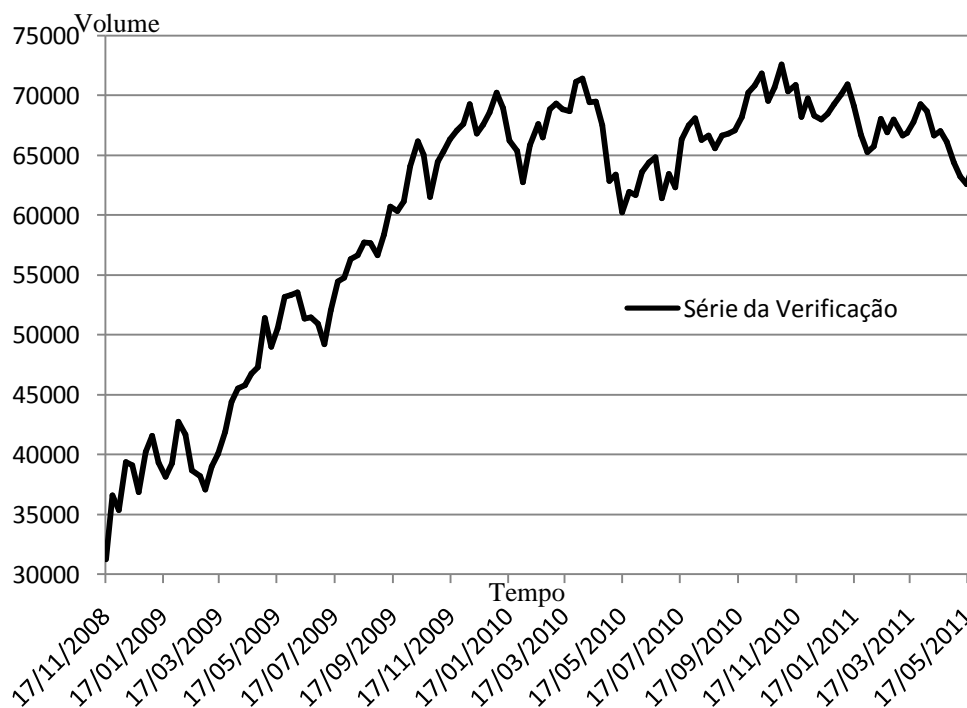


Figura 25 – Série de Dados da Verificação

Fonte: Organizada pelos autores.

Para a análise dos resultados obtidos na previsão do retorno do volume utilizou-se três critérios sendo estes descritos nos subcapítulos 2.5.6.2 e 2.5.6.4. Entretanto, dois destes critérios, servirão para identificar qual modelo tem menor erro. Por outro lado, o ultimo critério a ser analisado será o mais conclusivo dos três, pois verificará qual tem o melhor ajuste dos parâmetros de suavização nos modelos para os dados utilizados. O primeiro critério a ser analisando será o Erro Médio Global, que identificará qual rede neural obteve menor erro. Em seguida, o segundo critério a ser examinado será o Erro Médio Quadrático, com a finalidade de confirmar os resultados da comparação anterior. E o ultimo critério, fechando o ciclo de comparações, constatará qual das duas redes neurais tem o maior poder de explicação das variáveis independentes com as variáveis dependentes.

A Tabela 12 demonstra a comparação entre os dois modelos seguindo os critérios citados anteriormente. As redes neurais utilizadas foram à rede MLP de aprendizado supervisionado com os parâmetros testados no subcapítulo 4.1 e a rede neural **Kohonen** de aprendizado não supervisionado empregando os parâmetros avaliados no subcapítulo 4.2.

Tabela 12 - MLP VS **Kohonen**

	Rede Neural MLP	Rede Neural Kohonen
EMG	0,107997	1,42282
EMQ	0,060241	0,79365
R ² Ajustado	0,000213	0,0000146

Fonte: Organizada pelos autores.

Como demonstrado na Tabela 12, segundo os três critérios de avaliação utilizados não se pode relacionar previsões satisfatórias de oscilações provenientes de uma crise econômica, demonstrando que os métodos, etapas e séries de dados descritos nos subcapítulos anteriores utilizados não auxiliam na previsão de crises. Entretanto, é indispensável acrescentar que mesmo não obtendo resultados satisfatórios quanto à previsão de crises econômicas pode-se comparar a previsão das duas redes neurais corroborando um dos objetivos do trabalho. Na Tabela 12, como mencionado anteriormente apesar dos resultados não serem positivos, podemos observar que a rede neural MLP obteve melhores resultados na previsão de oscilações do volume que a rede neural **Kohonen** em relação ao mercado acionário em épocas de crises financeiras.

Na analisando dos resultados das três equações dos dois modelos na Tabela 12, pode-se perceber que a rede neural MLP obteve um melhor desempenho nos três critérios de avaliação adotados neste trabalho. Tendo-se em mente que o desempenho apresentado pelo EMQ e EMG das duas redes neurais obtiveram um acréscimo significativo em relação à validação da verificação cruzada, evidenciando que os métodos, etapas e séries de dados descritos no subcapítulo 3.4, não são possíveis para previsão de oscilações provenientes de uma crise econômica.

Junte-se a isso o fato de que o coeficiente de determinação ajustado descrito na Equação (2.11) da rede neural MLP demonstrou uma superação do poder explicativo da variável X exercendo em relação à variável Y , do que a rede neural **Kohonen**, corroborando com os estudos de Campos, Coelho, Maia, (2008), que compararam modelos econométricos, redes neuro-fuzzy e rede neural MLP, Carron et al. (2012), que previu mudanças climáticas com uma rede neural, Huang et al. (2012), que utilizou um modelo de rede neurais com um algoritmo híbrido para a previsão do índice de fusão (MI) para determinar a qualidade e o grau de polimerização dos produtos que utilizam polipropileno em sua composição e Al-Jumaily e Chen (2012) que estimou a rigidez do músculo das vias aéreas utilizando redes neural MLP.

Entretanto, apesar dos resultados satisfatórios na comparação das duas redes neurais, para que a modelagem do modelo vencedor esteja de acordo com estudos empíricos realizados é preciso realizar os testes propostos no subcapítulo 3.5. Estes testes realizados deverão corroborar com os resultados demonstrando que o modelo vencedor possui resíduo com a variância constante ou homocedastica, sem autocorrelação nas variáveis e com a contribuição positivos próximo a 100% entre as variáveis dependentes em relação as variáveis independentes.

4.4 Testes Realizados

Nesta etapa do trabalho será apresentada a análise dos testes realizados com a rede neural que obteve menor erro na etapa da verificação. Para esta análise foi utilizado o software R versão 2.13.2 com os pacotes necessários dos dois testes utilizados: o teste Breusch-Pagan (BREUSCH e PAGAN, 1979); White (WHITE, 1980); e o teste Durbin e Watson (1950). Para a análise da seleção de variáveis e eliminação de variáveis dependentes com pouca ou nenhuma explicação entre as variáveis independentes foi utilizado o valor das equações de Garson (1991) e Goh (1995) anexadas ao algoritmo backpropagation.

Como mencionado anteriormente, para verificar a heterocedasticidade dos resíduos da amostra foi desenvolvido o teste White. Com a análise dos resultados deste testes pode-se perceber que as três variáveis independentes tiveram um resíduo com a variância constante, ou seja, homocedásticos. Para esta análise foi levado em consideração os resultados obtidos do R^2 de 0,02163 correspondente a 4,9749 a baixo de 11,0705, valor crítico do qui-quadrado no nível de significância de 5%, não rejeitando H_0 , com os resíduos homocedásticos. Para confirma os resultados obtidos do teste White (1980), foi utilizado o teste Breusch-Pagan que obteve o valor do p-value de 0,06928 superior a 0,05, não rejeitando H_0 , com os resíduos homocedásticos, confirmando o resultado do teste White (1980) para resíduos do modelo homocedásticos.

Como mencionado anteriormente o ultimo teste a ser analisado, foi o teste de Durbin-Watson, que avaliou a autocorrelação das variáveis, ou seja, em uma previsão pressupõe-se que o termo de erro de qualquer variável não e influenciada e nem influencia pelo termo de erro de suas consequintes. Para esta análise foi levado em consideração os resultados das três variáveis utilizadas na rede neural MLP, volume, tempo e constante. Como demonstrado no resultado do *p-value* de 0,04953, a baixo de 0,05 rejeitando H_0 para resíduos das variáveis não autocorrelacionadas. Então, seguindo o valor de significância de 0,05 para

uma entrada com três variáveis, o modelo referido obtém um valor da estatística d do teste Durbin-Watson com 200 observações acima do valor crítico superior de 1,799 com evidências conclusivas em relação a não presenças de correlação serial positiva de primeira ordem.

A inclusão de entradas irrelevantes pode prejudicar o processo de treinamento, tendo por resultado, correlações indevidas nos dados que estão sendo modelados (VALENÇA e LUDERMIR, 2007). Para eliminar o problema da “caixa preta” da rede neural utilizada, será desenvolvida uma análise dos resultados da Equação (3.3 b), testando a relação das variáveis dependentes com as variáveis independentes, Tabela 13, a seguir. Logo a metodologia proposta vem ajudar a compreender a analogia e a importância das variáveis de entrada como também permitir uma eliminação de variáveis não necessárias para representar o processo que esta sendo modelado. Embora a rede neural utilize três variáveis como entradas somente foram testadas as variáveis do volume e do tempo, pois a variável constante tem a função de segurar a função de normalização, ou seja, não podendo ser eliminada.

Tabela 13 - Teste da “Caixa Preta”

	<i>Cont_j</i> (%)
Volume	66,3
Tempo	31,7

Fonte: Organizada pelos autores.

Na Tabela 13, pode-se observar que as duas variáveis utilizadas na rede neural obtiveram resultados de contribuição positivos, concluindo-se que em relação ao processo de modelagem todas as variáveis são fundamentais para o resultado do treinamento. Como decorrência disso, pode-se afirmar que a rede neural MLP modelada neste trabalho não tem o problema da “caixa preta” em suas variáveis de entrada.

4.5 Limitações

Os diferentes métodos da inteligência computacional podem ser aplicados de maneiras variadas. Como vimos anteriormente o termo *neural network*, por exemplo, abrange diferentes algoritmos de treinamento e diversas configurações possíveis para uma rede neural. O estudo realizado apresentou limitações importantes quanto a pouco investimento intelectual atualmente da comunidade acadêmica em relação a redes com aprendizados não supervisionados, pois redes com aprendizados supervisionados na maioria dos casos proporcionam resultados mais explicativos e conclusivos no conjunto das etapas do

treinamento. Entretanto, em alguns casos esta relação não explicativa entre variáveis de entrada e o treinamento de uma rede neural **Kohonen** pode trazer benefícios em vários casos com variáveis não temporais segundo **Kohonen** (1982), Milioni et al. (2007), Khashei et al. (2012) e Anitha et al. (2012).

Destacam-se ainda como limitações do trabalho a pouca utilização de estudos de caso, com competição entre séries e modelos de redes neurais com aprendizados distintos. Outra limitação do trabalho refere-se a poucos investimentos financeiros e intelectuais nas áreas de previsão de crises econômicas. Entretanto, a maior limitação deste trabalho foi pelo pouco tempo de elaboração e treinamento das redes neurais.

5 CONSIDERAÇÕES FINAIS

Como mencionado anteriormente, o mercado acionário replica a determinação de milhões de pessoas (HISSA, 2008). A partir desta constatação, não há como antecipar minuciosamente o futuro dos ativos ou de uma carteira de ações utilizando uma previsão simples. Entretanto, uma previsão idealista, ou seja, aquela que se utiliza de todos os fatores internos e externos de um projeto financeiro, como variável de entrada do modelo, seria possível até mesmo a previsão de uma crise econômica.

Então, resta afirmar que, para a previsão de uma crise econômica é necessário a utilização, como variáveis de entrada, de todos os fatores e decisões que possam atingir o projeto financeiro, ou seja, a utilização de uma previsão ideal. Pois, não há dúvidas de que a utilização de uma previsão idealista no mercado financeiro para o gerenciamento de risco eliminaria a zero ou muito próximo de zero o risco ao maximizar o retorno. Entretanto, a utilização de múltiplas variáveis tornaria o projeto inviável maximizando o problema da “caixa preta” dos métodos.

Porém, segundo Faraway e Chatfield (1998), a previsão idealista de um projeto depende de mais três critérios além do citado anteriormente: dos tipos de dados como variáveis apropriadas; da técnica do analista na escolha do modelo; e principalmente dos métodos numéricos usados para ajustar o modelo e compor a saída desejada. Como decorrência disso, o modelo utilizado tanto quanto principalmente o treinamento do mesmo é importante para uma previsão bem sucedida.

Entre modelos pode-se afirmar, segundo Milioni et al. (2007), Sobreiro et al. (2009), Ceretta et al. (2010) e Galão et al. (2011) que as redes neurais são adequadas para a previsão de séries temporais. Então, com base na eliminação dos dois primeiros critérios de Faraway e Chatfield (1998), pelas afirmações dos autores citados no parágrafo anterior propõem-se a análise do último critério, ou seja, o ajuste e composição da saída desejada. Quanto ao ajuste dos pesos dos modelos, pode-se afirmar no capítulo 4, que foram bem modelados. Entretanto, a responsabilidade de uma previsão bem sucedida volta-se a composição da saída desejada da rede neural.

Então, para contemplar o objeto deste estudo, objetivou-se formular três grandes objetivos principais resumidos em desenvolver subcapítulo 2.5.4 e 2.5.5, treinar subcapítulo 3.4 e validar subcapítulo 4.3 dois métodos heurísticos para antever oscilações provenientes de uma crise econômica, sem problema da “caixa preta” dos métodos. No término de todas as

etapas descritas anteriormente, concluiu-se, enfim, que apesar de alguns modelos de redes neurais obterem resultados satisfatórios na previsão do mercado acionário como demonstrado em Ceretta et al. (2010) e Galão et al. (2011). Ao utilizarmos uma análise técnica como a teoria das ondas de Elliott na saída desejada da rede neural não foi possível observar os mesmo resultados positivos da previsão do mercado acionário quanto à previsão de oscilações provenientes de uma crise econômica.

Ainda em relação aos resultados da previsão de oscilações de crises econômicas subcapítulo 4.3. Conclui-se, que a periodicidade dos dados interferiu negativamente nos resultados do trabalho. Pois, modelos para a previsão de crises econômicas que utilizam dados muito esparsos como semanais e a teoria das ondas de Elliott para compor a saída desejada das redes neurais, não são eficientes, pois séries de dados semanais eliminam a precisão da contagem de ondas.

A segunda conclusão elaborada em relação aos resultados não positivos, diz respeito à utilização da teoria das ondas de Elliott como variável de entrada. Pois, podemos observar no subcapítulo 2.6, que Elliott em sua teoria empírica, utilizou um método científico como a seqüência de Fibonacci para transformá-la em teórica. Entretanto, como demonstrado nos resultados deste trabalho, a analogia feita para a criação da teoria das ondas de Elliott entre a teoria do comportamento psicológico das massas e a seqüência de Fibonacci, se mostrou incapaz de prever oscilações do mercado em uma série correspondente a uma crise econômica. Enfim, segundo as afirmações anteriores, conclui-se, que nos dias atuais, uma previsão idealista do mercado acionário utilizando a teoria das ondas de Elliott e series temporais, esta fora da realidade. Porém, uma previsão simples pode ser utilizada em vários casos com resultados satisfatórios como demonstrado em Milioni et al. (2007), Sobreiro et al. (2009), Ceretta et al. (2010) e Galão et al. (2011).

Apesar dos resultados negativos na previsão, com a análise da etapa da verificação subcapítulo 4.3, foi possível comparar os resultados da previsão das duas redes neurais chegando a conclusões satisfatórias nesta parte do trabalho. Na comparação da previsão dos dois modelos de rede neural, foram avaliados três critérios que são: o menor erro médio global; o menor erro médio quadrático; e o maior R^2 ajustado. O primeiro critério levou em consideração o menor valor do erro médio global obtido entre a rede neural MLP e **Kohonen**. Nesta comparação pode-se perceber que o EMG da rede neural **Kohonen** obteve um valor relativamente mais alto do que a rede MLP, fazendo assim, com que o primeiro critério de validação seja a favor da rede neural MLP. O segundo critério de avaliação foi o erro médio quadrático dos modelos, onde se pode perceber que a rede neural MLP também obteve um

melhor desempenho levando-se em consideração o menor valor com 0,060241 que a rede **Kohonen** com 0,79365. O terceiro método de avaliação foi à comparação do poder de explicação entre as variáveis dos modelos, onde se pode perceber que a rede neural MLP foi superior obtendo um resultado de 0,000213 do R^2 ajustado contra 0,0000146 da rede **Kohonen**.

Concluiu-se, portanto, na comparação dos métodos um dos objetivos do trabalho, que redes neurais com algoritmos de aprendizado não supervisionado e séries temporais, neste caso eliminam um processo manual no treinamento da rede neural, ocasionou uma falha na previsão no conjunto treinamento e validação. Pois, ao eliminar o processo do supervisor, elimina-se também uma etapa de controle dos resultados do treinamento da rede, ocasionando uma falha na busca do objetivo pelas quais foi treinada. Esta falha desvia os resultados do treinamento calculado pela rede neural não supervisionada, do objetivo inicial, treinando-a para outro objetivo qualquer.

Então, ao se utilizar uma rede neural **Kohonen** e um serie temporal de dados como o retorno do volume do Ibovespa, pode ocasionar falhas na procura por um padrão parcialmente aleatório pela rede neural **Kohonen**. Pois, a partir da escolha dos valores de parâmetros, não há como saber que padrões a rede neural **Kohonen** baseou-se para o seu treinamento, deixando seus resultados ainda mais imprevisíveis na verificação. Esta constatação baseou-se principalmente nos resultados da comparação do conjunto treinamento e validação das duas redes neurais, onde se pode perceber que a rede MLP obteve resultados mais harmônicos que a rede neural **Kohonen**. Pois, os padrões que a rede MLP utilizou em todo seu treinamento foi o mesmo passado pelo supervisor. Entretanto, foram utilizados padrões aleatórios para o treinamento da rede neural **Kohonen**, levando-a a resultados mais esparsos e menos satisfatórios, entre o treinamento/aprendizado e a verificação. Os resultados indicaram ainda, que quanto menor é à periodicidade dos dados de entrada, maior é a capacidade do modelo em realizar boas previsões corroborando com Coelho et al. (2008), pois quanto menor for a periodicidade dos dados menor será o erro.

5.1 Sugestões para Trabalhos Futuros

Apesar de obter resultados não satisfatórios quanto à previsão de oscilações provenientes de crises econômicas. Este trabalho abriu uma grande variedade de opções disponíveis para serem utilizadas em trabalhos futuros, como a utilização do volume ou preço diário, minuto a minuto e ou segundo a segundo. Esta medida eliminaria o problema da

periodicidade dos dados encontrado nos resultados deste trabalho, eliminado também alguns processos de identificação de ondas.

Acrescente-se ainda, para sugestões de trabalhos futuros, a utilização de varias outras teorias de previsão proveniente da análise técnica e análise fundamentalista. Esta medida traria mais alguns pontos negativos e ou positivos sobre a utilização da analise técnica no treinamento de uma rede neural, podendo assim, haver algumas comparações dos resultados deste trabalho com trabalhos futuros. Ainda, quanto aos resultados negativos, sugere se como trabalhos futuros utilizar somente a relação do mercado acionário com a sequência de Fibonacci, não levando em consideração a teoria das ondas de Elliott.

Também, sugere-se desenvolver uma técnica ou utilizar uma já existente que identifique o padrão utilizado pela rede neural **Kohonen**, e a partir disto, treinar a rede neural MLP com o mesmo padrão. Em consequência disso, os resultados das duas redes transformam-se em correlacionados aos mesmos padrões proposto as duas redes neurais, não eliminando um processo manual no treinamento da rede neural **Kohonen**.

Em relação à comparação das redes neurais, sugere-se para trabalhos futuros a utilização de outras redes não supervisionadas como as redes **GMDH** e redes supervisionadas como as redes *Radial Basic Functions*, redes *Non-Linear Signoidal Regression Blocks* e modelos híbridos de redes com outros métodos heurísticos.

REFERÊNCIAS

- AIUBE, F. A. L.. **Econometria para Séries Financeiras**. Pontifícia Universidade Católica do Rio de Janeiro (Trabalho de conclusão de curso de Engenharia Industrial), Rio de Janeiro, 2007.
- AL-JUMAILY, A.; CHEN, L.. Estimation of airway smooth muscle stiffness changes due to length oscillation using artificial neural network. **Journal of Theoretical Biolog.** v.310, p.115 n.4. New Zealand, 2012.
- ALVES, J. S. Análise comparativa e teste empírico da validade dos modelos CAPM tradicional e condicional: o caso das ações da Petrobrás. **Revista Ciência Admin.**, Fortaleza, v. 13, n. 1, p. 147-157, 2007.
- ANITHA, J.; VIJILA, C. K. S.; SELVAKIMAR, A. I.; INDUMATHY. A.; HEMANTH, D. J.. Automatéd multi-level pathology identification techniques for abnormal retinal images using artificial neural networks. **Journal BJO: Clinical Science**, v. 96, n. 10, p. 220-223, 2012.
- BARBA, F. G. **Impactos da crise de 2007/2008 nos mercados de capitais latino-americanos**. Universidade Federal de Santa Maria (Dissertação de Mestrado), Santa Maria, 2011.
- BARBOSA, A. H.; FREITAS, M. S. R.; NEVES, F. A. Confiabilidade Estrutural Utilizando o Método de Monte Carlo e Redes Neurais. **REM: Revista Escola de Minas**. Minas Gerais, v.58, n. 3, p. 247-255, 2005.
- BARTRAM, S.; BODNAR, G. No place to hide: the global crisis in equity markets in 2008 / 2009. **Journal of International Money and Finance**, v. 28, p. 1246-1292, 2009.
- BLANCHARD, O.. The Crisis: Basic Mechanisms, and Appropriate Policies. **Working Paper**. Cambridge, 2008.
- BM&FBOVESPA: **A nova bolsa. Sobre a bolsa**. São Paulo. 1968. Disponível em < <http://www.bmfbovespa.com.br/pt-br/intros/intro-sobre-a-bolsa.aspx?idioma=pt-br> > Acesso em 15 de fevereiro de 2012.
- BRESSAN, A. A. Tomada de decisão em futuros agropecuários com modelos de previsão de séries temporais. **RAE-eletrônica**. São Paulo, v. 3, n. 1, Art. 9, 2004.
- BREUSH, T.; PAGAN, A.. A Simple Test for Heteroscedasticity and Random Coefficient Variation. **Journal The Econometric Society**, v. 7, n. 5, p.1.287-1.294, 1979.
- CAMPOS, R. J. **Previsão de séries temporais com aplicações a séries de consumo de energia elétrica**. Universidade Federal de Minas Gerais (Dissertação de Mestrado). Belo Horizonte, 2008.
- CARGNELUTTI, A. F.; MATZENAUER, R.; RADIN, B.; MALUF, J. R. T.; HAUSER, L. Dimensionamento da amostra para a estimação da média de precipitação pluvial mensal em

diferentes locais do Estado do Rio Grande do Sul. **Revista de Ciências Rurais**. vol.40 no.1 Santa Maria, 2009.

CARRON, A. V.; DUTOT, A. L.; LOMBARDO, T.; CHABAS, A.. Predicting changes of glass optical properties in polluted atmospheric environment by a neural network model. **Journal Atmospheric Environment**. v.54, p.141-148, 2012.

CAVALCANTE, F.. MISUMI, J. Y.; RUDGE, L. F. **Mercado financeiro e de capitais o que é e como funciona**. ed. 6. Rio de Janeiro: Elsevier, 2005.

CERETTA, P. S.; RIGHI, M. B; SCHLENDER, S. G. Previsão do preço do soja: uma comparação entre os modelos ARIMA e redes neurais artificiais. **Revista Informações Econômicas**, Instituto de Economia Agrícola - Centro de Comunicação e Transferência do Conhecimento. São Paulo. v. 40, n. 9, 2010.

CHAVES, Evenice Santos. Nina Rodrigues: sua interpretação do evolucionismo social e da psicologia das massas nos primórdios da psicologia social brasileira. **Revista Psicol. estud.** [online]. vol.8, n.2, pp. 29-37, 2003.

COELHO, L. S.; SANTOS, A. A. P.; COSTA JR, N. C. A. Podemos prever a taxa de cambio brasileira? Evidencias empírica utilizando inteligência computacional e modelos econométricos. **Revista Gestão de Produção**. São Carlos, v. 15, n. 3, 2008.

COSTA, J. A. F.; ANDRADE NETTO, M. L.. Segmentação de mapas auto-organizáveis com espaço de saída 3-D. **Sba Controle & Automação**, Natal, v. 18, n. 2, jun, 2007 . Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-17592007000200002&lng=pt&nrm=iso>. acessos em 3 Jan. 2012. <http://dx.doi.org/10.1590/S0103-17592007000200002>.

CUNHA, G. B.; MOURA, J. F. L.; NAVES, E. L. M.; ANDRADE, A. O.; PEREIRA, A. A.; MILAGRE, S. T.. A utilização de uma Rede Neural Artificial para Previsão de Incidência da Malária no Município de Cantá, Estado de Roraima. **Revista da Sociedade Brasileira de Medicina Tropical (on-line)**. Uberaba-MG. set./out., 2010.

DEBASTIANI, C. A. **Análise Técnica de Ações: Identificando Oportunidades de Compra e Venda**. São Paulo: Novatéc, 2008.

DIDIER, T.; LOVE, I.; PERÍA, M. S. What explains stock market's vulnerability to the 2007-2008 crisis?. **World Bank policy research working paper**, 2010.

DOUADY, S.; COUDER, Y.. Phyllotaxis as a Dynamical Self Organizing Process Part I: The Spiral Modes Resulting from Time-Periodic Iterations. **J. Theor. Biol. - Journal: Theory Biology**, France, v. 067, p. 144-163, 1995.

DURBIN, J.; WATSON, G. S.. Testing for Serial Correlation in Least Squares Regression. **Journal of Biometrika Trust**. v. 38, p. 159-171, 1950.

ELLIOTT, R. N.. **Nature'S Law: The Secret Of The Universe**. England: Snowball, 2011.

FARAWAY, J.; CHATFIELD, C. Time series forecasting with neural networks: a case study. **Royal statistical society journal**. v. 47, n. 2, p. 231-250, 1998.

FARIAS, H. P.; SÁFADI, T.. Causalidade Entre as Principais Bolsas de Valores do Mundo. **RAM - Revista de Administração Mackenzie**, São Paulo, v. 11, n. 2, 2010.

FERNANDES, L. G. L., PORTUGAL, M. S., NAVAUUX, P. O. A.. **Previsão de séries de tempo: redes neurais artificiais e modelos estruturais**. In: 17 ENCONTRO BRASILEIRO DE ECONOMETRIA.v. 17, p. 399-416. Salvador: Soc. Bras. de Econometria, 1995.

FARHI, M.. O impacto dos ciclos de liquidez no Brasil: Mercados financeiros, taxa de câmbio, preços e política monetária. **Revista Política Econômica em Foco**. Campinas – SP, n. 7. 2006.

GALÃO, O. F.; BORSATO, D.; PINTO, J. P.; VISENTAINER, J. V.; PANIZZI, Mercedes C. C. . Artificial Neural Network in the Classification and Identification of Soybean Cultivars by Planting Region. **Journal of the Brazilian Chemical Society**. v. 22, n. 1, p. 142-147, 2011.

GARSON, G. D.. Interpreting neural-network connection weights. **Artif. Intell. Expert**. San Francisco, v. 6, n. 4, p. 46-51, 1991.

GRAHAM, B.; DODD, D. L.. **Security Analysis: Sixth Edition, Foreword by Warren Buffett**. 6. ed. New York: McGraw-Hill, 2009.

GOH, A. T. C. Backpropagation neural network for modeling complex systems. **Artificial Intelligence in Engineering**, v. 9, n. 3, p. 143-15. Deutsch, 1995.

GOMES, G. S. da S.; LUDERMIR, T. B.. REDES NEURAIS ARTIFICIAIS COM FUNÇÕES DE ATIVAÇÃO COMPLEMENTO LOGLOG E PROBIT PARA APROXIMAR FUNÇÕES NA PRESENÇA DE OBSERVAÇÕES EXTREMAS. **Revista da Sociedade Brasileira de Redes Neurais (SBRN)**, v. 6, n. 2, p. 142- 153. Rio de Janeiro, 2008.

GONÇALVES, E. B.. **Análise de Risco de Credito com o Uso de Modelo de Regressão Logística, Redes Neurais Artificiais e Algoritmos Genéticos**. Dissertação (Mestrado em Administração) - Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 2005.

GUIMARÃES, A. M.. Modulo de Validação Cruzada para Treinamento de Redes Neurais Artificiais com Algoritmos Backpropagation e Resilient Propagation. **Revista Publicatio UEPG**, Ponta Grossa, v. 14, n. 01, 2008.

HAYASHI, A. D.. **APLICAÇÃO DOS FRACTAIS AO MERCADO DE CAPITAIS UTILIZANDO-SE AS ELLIOTT WAVES**. Universidade Federal de Santa Catarina (Dissertação de Mestrado), Florianópolis, 2002.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. 3. ed, São Paulo: Bookman, 2001.

HASSAN, R. M. D.; NATH B. **Stock Market Forecasting Using Hidden Markov Model: A New Approach**. In: Proceedings in the 5th International Conference on Intelligent Systems Design and Applications (ISDA'05). IEEE Computer Society, 2005.

HISSA, M. B. **Sobreviva na Bolsa de Valores: Um roteiro para acumular capital com ações e renda variável**. 3. ed.. Rio de Janeiro: Elsevier, 2008.

HUANG, M.; LIU, X.; LI, J.. Melt Index Prediction by RBF Neural Network with an ICO-VSA Hybrid Optimization Algorithm. **Journal of Applied Polymer Science**. v.126, p.519-526, 2012.

HUYGHEBAERT, N.; WANG, L. The co-movement of stock markets in East Asia: did the 1997-1998 Asian financial crisis really strengthen stock market integration. **Rev. China Economic**, v. 21, p. 98-112, 2010.

JONES J.; WILSON, W.. **An Incomplete Education**. 1. ed., Unknown, Ballantine Books, p. 699, 2006.

JUDD J. S, **Neural Network design and Complexity of Learning**. MIT Press, USA, 1990.

KATAGIRI, H.; NISHIZAKI, I.; HAYASHIDA, T.; KADOMA, T.; Multiobjective Evolutionary Optimization of Training and Topology of Recurrent Neural Networks for Time-Series Prediction. **The Computer Journal**, v. 55, n. 3. England: London, 2012.

KHALID, A.M.; RAJAGURU, G. Financial market contagion: evidence from Asian crisis using multivariate GARCH approach. **Bond University working paper**. 2007.

KHASHEI, M.; BIJARI, M.; ARDALI G. A. R. Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks (PNNs). **Journal Elsevier: Computers & Industrial Engineering**, v. 63, n. 1, p. 37-45, 2012.

KIM, K.; LEE, W. B. Stock market prediction using artificial neural networks with optimal feature transformation. *Neural Computing & Applications*. **Journal: Springer-Verlag**, v. 13, pp. 255-260. Milão: Heidelberg, 2004.

KIMOTO, T. K.; ASAKAWA, M; YODA, M. T. Stock market prediction system with modular neural networks. In: **Proceedings of the International Joint Conference on Neural Networks**, v. 1, pp. 1-6. New York, 1990.

KINDLEBERGER, C. P.; ALIBER, R. Z.. **Manias, Panics and Crashes: a history of financial crises**. 5. ed. New York: Palgrave MacMillan, 2005.

KOHONEN, T. Self-organization formation of topologically correct feature maps. **Biological Cybernetics**. v. 43, pp. 59-69, 1982.

LAEVEN, L.; VALENCIA, F. Systemic Banking Crises: A New Database. **IMF Working Paper**. WP/08/224. N.W., Washington, D.C., 2008.

LAWRENCE R. **Forecasting stock prices using neural networks**, Department of Computer Science. University of Manitobo. December, 1997.

LI, X.; FU, X. Lag synchronization of chaotic delayed neural networks via impulsive control. **IMA Journal of Mathematical Control and Information**. v. 29. p. 133-145 Reino Unido: Oxford, 2012.

LOPES, L. F. D et al.,. **Caderno Didático: estatística geral**. Santa Maria: UFSM, 3. ed., CCNE, 2008.

LUDWIG J. R.; O.; MONTGOMERY, E. **Redes Neurais: Fundamentos e Aplicações com Programa em C**. 1. ed. Rio de Janeiro: Ciência Moderna, 2007.

LUQUET, M. **Guia valor econômico de finanças pessoais**. São Paulo: Globo, 2000.

McCULLOCH, W.; PITTS, W. A logical calculus of the ideas immanent in neurons activity. Springer New York, **Journal: Bulletin of Mathematical Biophysics**, v. 5, n. 4, 1943.

MELLO, P.; SPOLADOR, H. **Crises financeiras: uma história de quebras, pânico e especulações do mercado**. São Paulo: Saint Paul Editora, 2007.

MILIONI, A. Z.; MELO, B.; NASCIMENTO, C. L. J.. Daily and Monthly Sugar Price Forecasting Using The Mixture of Local Expert Models. **Revista: Pesquisa Operacional**. v.27, n.2, p.235-246. Rio de Janeiro, 2007.

NASCIMENTO J. R. L. C.; YONEYANA, T. **Inteligência Artificial em Controle de Automação**. São Paulo: FAPESP, 2004.

NORONHA, M.. **Análise Técnica: Teorias Ferramentas Estratégias**. 7ª Edição. Rio de Janeiro, EDITEC, 2009.

OLIVEIRA, A. C. S; SOUZA, A. A; LACERDA, W. S.. Aplicação de Redes Neurais Artificiais na Previsão da Produção de Alcool. **Revista Ciência e Agrotecnologia**, Lavras-MG. v. 34, n. 2, p. 279-284, mar./abr., 2010.

OLIVEIRA, J. N. J.; CASTELAR, I.; TATIWA, R.. **Previsão da Taxa de Crescimento do Produto Agrícola Brasileiro: Uma Aplicação de Modelos de Índice de Difusão Linear e Não Linear**. In: XI Encontro Regional de Economia da ANPEC. Anais do XI Encontro Regional de Economia da ANPEC. Fortaleza, 2006.

PANDORFI, H.; SILVA, I. J. O.; SARNIGHAUSEN, V. C. R.; VIEIRA, F. M. C.; NASCIMENTO, S. T.; GUISELINI, C.. Uso de Redes Neurais Artificiais para Predição de Índices Zootécnicos nas Fases de Gestação e Maternidade na Suinocultura. **Revista Brasileira de Zootecnia**. Pernambuco - PE. v. 40, n. 3, p. 676-681, 2011.

PHUA, P.K.H.; MING, D.; LIN, W. **Neural Network With Genetic Algorithms For Stocks Prediction**. Fifth Conference of the Association of Asian-Pacific Operations Research Societies, 5th - 7th July. Singapore, 2000.

RIGOBON, R. **International financial contagion: theory and evidence in evolution**. Virginia, EUA: The research foundation of AIMR, 2002.

ROBERT, D. E.; MAGEE, J.; BASSETTI, W. H. C. (Editor). **Technical Analysis of Stock Trends**. American Management Association, 2007.

SAHOO, A. K.; ZUO, M. J.; TIWARI, M. K.. A data clustering algorithm for stratified data partitioning in artificial neural network. **Journal Elsevier: Expert Systems with Applications**, v. 39, n. 8, p. 7004–7014, 2012.

SCHALKOFF, R. J. Artificial neural networks. **SAGE Publication - Progress in Physical Geography**. New York: McGraw-Hill, 2001.

SMITH, K. A.; GUPTA, J. N. D. Neural networks in business: techniques and applications for the operations researcher. v. 27. p. 1023-1044. **Oxford: Computers & Operations Research**. Elsevier, 2000.

SOARES, S.; et al.. **A crise internacional e possíveis repercussões: primeiras análises**. Instituição de Pesquisas Econômicas Aplicadas: IPEA. São Paulo, 2009.

SOBREIRO, V. A.; ARAÚJO, P. H. S. L. NEGANO, M. S.. Precificação do Etanol Utilizando Técnicas de Redes Neurais Artificiais. **Revista de Administração**. São Paulo. v. 44, n. 1, 2009.

SOUSA, A. F. M.; VALENÇA, M.. **Seleção de Variáveis: Um Sistema Híbrido Baseado em Colônia de Formigas e Redes Neurais MLP**. Trabalho de Conclusão de Curso Engenharia da Computação Escola Politécnica de Pernambuco. Recife, 2010.

STEINER, M.. WITTKEMPER, Hans-Georg. Neural networks as an alternative stock market model. **Journal Neural Networks in the Capital Markets**, chapter 9, p 137–148. John Wiley and Sons, 1995.

STONE, O.; PRESSMAN, E. R.; KOPELOFF, E.; CAMON, A.; COSTA, C. D.; DOUGLAS, M. K.. **Wall Street: Money Never Sleeps**. [Filme-DVD]. Produção de Oliver Stone, Edward R. Pressman, Michael Kirk Douglas, Eric Kopeloff, Alessandro Camon, Celia D. Costas, direção de Oliver Stone. Nova York. Twentieth Century Fox, 2011. 1 DVD, 132 min. color. Son.

TORRES, E. T.;. **Entendo a Crise do Subprime**. Informativo do BNDS visão do desenvolvimento. São Paulo, 2008.

TORRES, E. T.; BORÇA JR., G. R.. **A Crise do Subprime ainda não acabou**. Informativo do BNDS visão do desenvolvimento. São Paulo, 2008.

VALENÇA, M. J. S.; LUDERMIR, T. B. Explicando a Relação entre as Variáveis de uma Rede Neural – Iluminando a “Caixa Preta. In: **XVII Simposio Brasileiro de Recursos Hídricos**. São Paulo, 2007.

VALENÇA, M. **Fundamento das Redes Neurais: Exemplos em Java**. 2 ed., Olinda: Livro Rápido – Elógica, 2009.

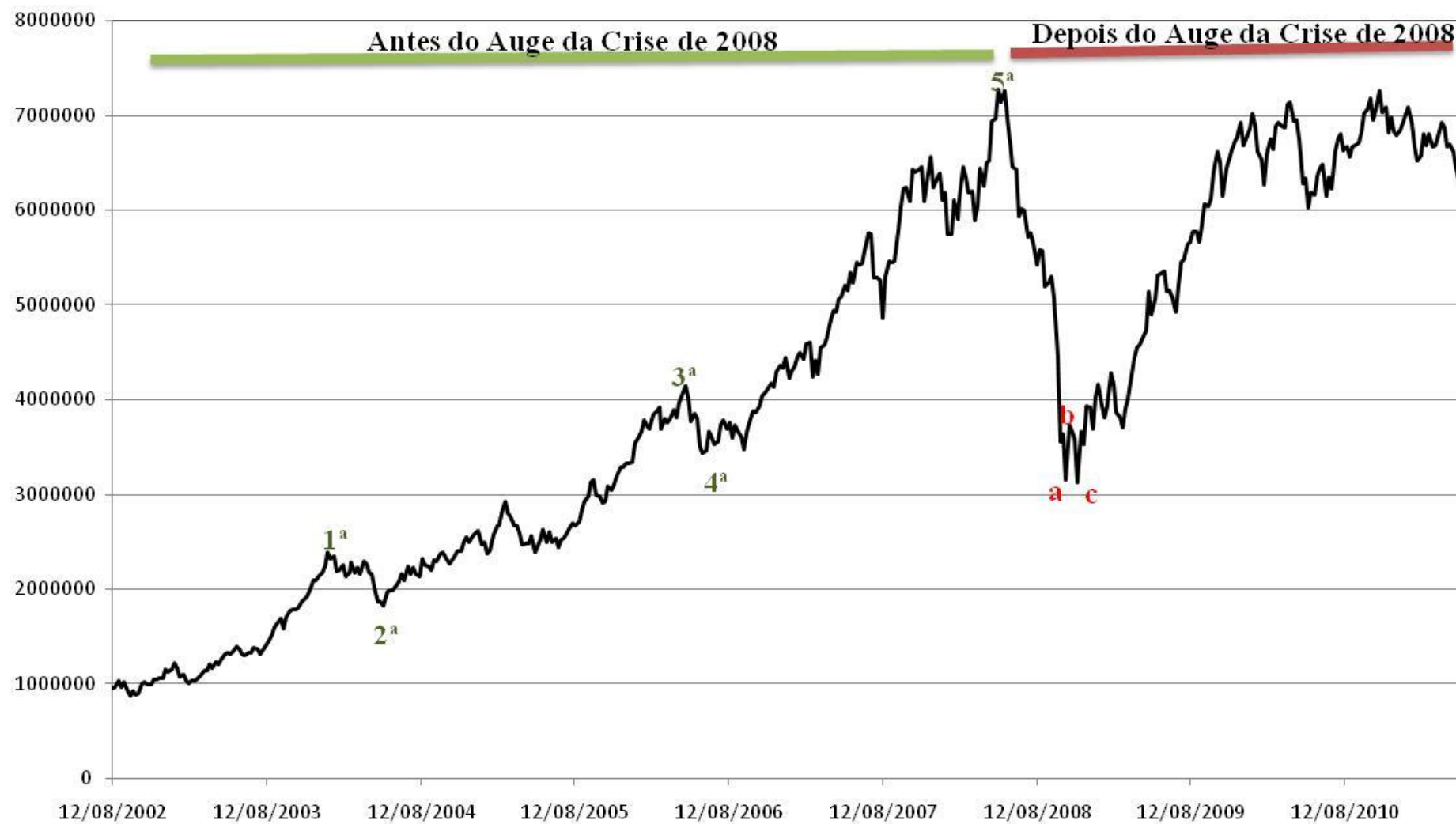
WEISZFLOG, W. **Dicionário Michaelis escolar língua portuguesa**. 3 ed., São Paulo: Melhoramentos, 2009.

WHITE, H. A heteroskedasticity-consistent covariance matrix estimation and a direct test for heteroskedasticity. **Journal of the Econometric Society**. v. 48, n. 4, p. 817-818, 1980.

WUERGUES, A. F. E; BORBA, J. A.. Redes Neurais, Lógica Nebulosa e Algoritmos Genéticos: Aplicações e Possibilidades em Finanças e Contabilidade. **Revista de Gestão da Tecnologia e Sistemas de Informação**. São Paulo. v. 9, n. 1, p.163-182, 2010.

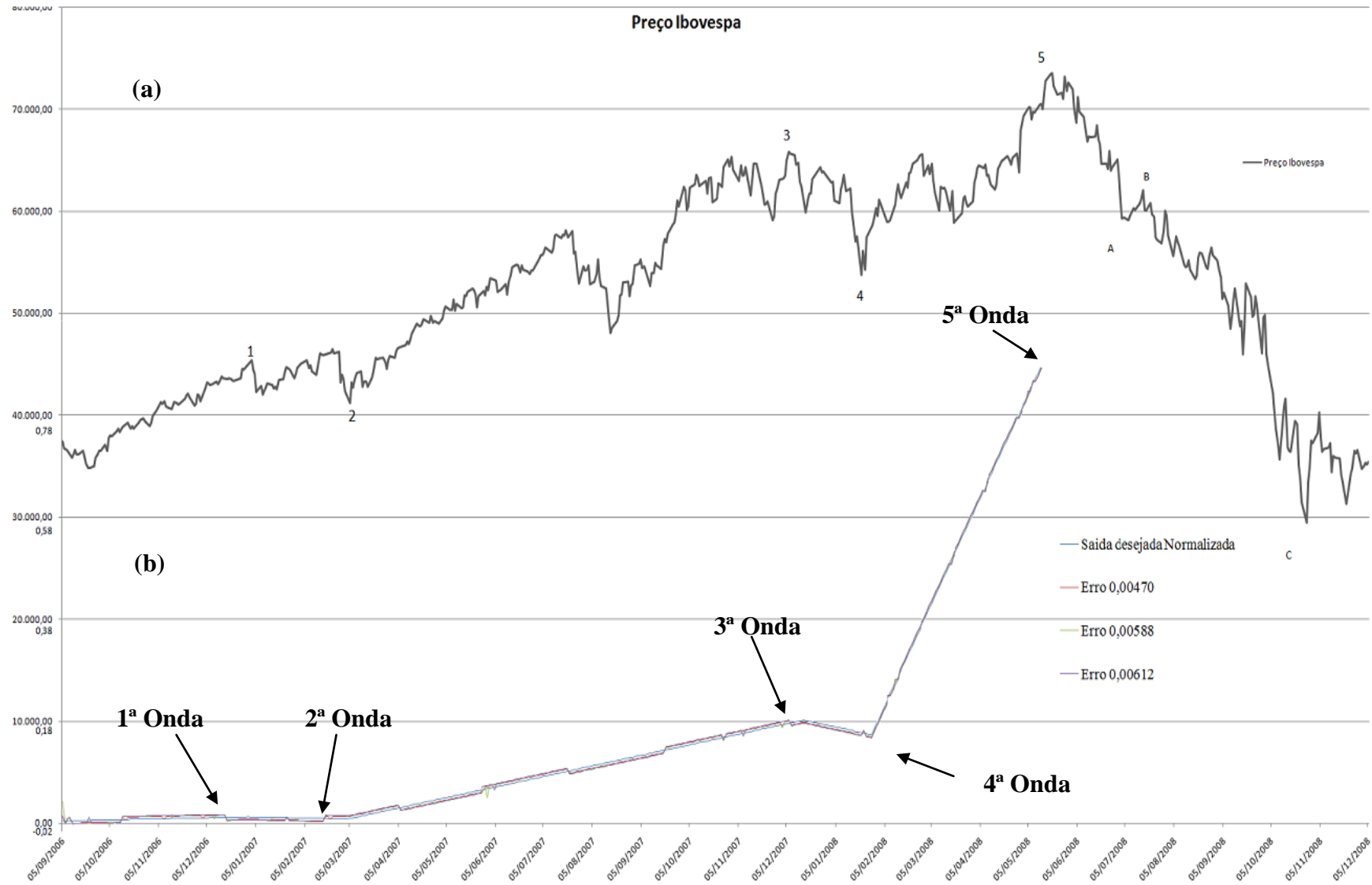
YONENAGA, H. W.; FIGUEIREDO S. R.. **Previsão do Preço da Soja Utilizando redes Neurais**. Trabalho de Conclusão de Curso (Programa de Pós-Graduação em Engenharia de Produção) - Universidade Federal de São Carlos – Rod, São Paulo, 1999.

ANEXOS

ANEXO A – Classificação dos dados de entrada conforme os padrões encontrados na Teoria das Ondas de Elliott

Fonte: Organizada pelos Autores

ANEXO B – Gráfico resultado do treinamento.



Fonte: Organizada Pelos autores.